Über die Schwierigkeit einfacher Additionen

Inaugural-Dissertation

zur Erlangung der Doktorwürde der Philosophischen Fakultät (I. Sektion)

vorgelegt von

Michael Braun

aus Pasing

München 1939

Über die Schwierigkeit einfacher Additionen

Inaugural-Dissertation

zur Erlangung der Doktorwürde der Philosophischen Fakultät (I. Sektion)

vorgelegt von

Michael Braun

aus Pasing

München 1939

Tag der mündlichen Prüfung:
28. Mai 1936

Referent: Geh. Rat Prof. Dr. Aloys Fischer

Gutachter: Professor Dr. Richard Pauli

R. UNIV.
BIBLIOTHEEK
LEIDEN

Meinen Eltern.

Inhalt.

Seite

Vorbemerkungen 5
Durchführung der Versuche 7
Versuchsmaterial 7
Versuchsanordnung 8
Einzelauswertung der beiden Versuchsgruppen 10
a) Fehlleistungsstatistik (1. Methode) 10
b) Zeitmeßmethode (2. Methode) 19
Vergleich der beiden Methoden 26
Schwierigkeitsfaktoren 30
Zusammenfassung 37

Abkürzungen:
Fehler = F Verbesserung = V Schwierigkeit = Schw

A. Vorbemerkungen.

Die Frage nach der Schwierigkeit einer Aufgabe scheint zunächst nur nach psychologischen Gesichtspunkten beantwortbar zu sein; denn ein Sachverhalt für sich kann zwar einfacher und verwickelter, aber nicht leichter und schwieriger sein. Dieselbe Arbeit kann je nach der Leistungsfähigkeit der betreffenden Person verschiedengroße Schwierigkeiten verursachen. Andererseits kann aber auch dieselbe Leistungsfähigkeit bei einer Person durch einen Wechsel in der Aufgabe verschieden stark beansprucht werden. Außerdem schließt das Auftreten einer Schwierigkeit stets nicht bloß eine individuelle Beanspruchung, sondern immer auch die tatsächliche Inangriffnahme bzw. Bewältigung einer objektiven Gegebenheit in sich. Es fragt sich nun, ob und wieweit die individuelle Leistungsfähigkeit einerseits bzw. die objektiven Gegebenheiten andererseits, die in der Struktur der jeweiligen Aufgabe liegen, die auftretende Schwierigkeit bestimmen.

Da die Aufgabe dieser Versuche die Untersuchung von objektiven Schwierigkeitsverhältnissen ist, soll die Verschiedenheit der individuellen Leistungsbeanspruchung möglichst gering sein, um damit ziemlich gleichmäßige Bedingungen zu schaffen für die großen Unterschiede hinsichtlich der individuellen Leistungsfähigkeit, d. h. es werden Additionen untersucht, die dem normalen erwachsenen Menschen keine nennenswerten Schwierigkeiten verursachen. Die Einschränkung auf die Verhältnisse bei den einfachen Additionen soll aber den Blick für die grundsätzliche Bedeutung der Fragestellung nicht trüben, sondern nur einen methodisch günstigen Angriffspunkt zu ihrer Lösung bieten. Es wird auch auf die Wichtigkeit dieser Frage für alle Zweige der angewandten Psychologie, die es mit der Schwierigkeit von Aufgaben zu tun haben, wie sie in der Testpsychologie, der Eignungsprüfung, der Lehrplantheorie oder der Psychiatrie praktisch auftreten können, hingewiesen.

Bisher sind in der Psychologie zur Prüfung von Leistungen vorwiegend intellektueller Art vor allem zwei Methoden benützt worden, die Fehlerstatistik und die Zeitmessung. Im ersteren Fall gilt als Schwierigkeitskriterium die relative Zahl von Falschlösungen einer Aufgabe bei gleichen Versuchsbedingungen für alle Aufgaben; bei der Zeitmessung gilt der Maßstab: Die Schwierigkeit wächst mit der steigenden Lösungsdauer einer Aufgabe. Außerdem fanden noch das Verdrängungsverfahren (22, 28) und die Urteilsstatistik (4) Verwendung: Hier gibt eine Anzahl von Personen an, welche Aufgaben besonders schwer bzw.

leicht zu lösen waren; diese subjektiven Urteile werden dann in einer Schwierigkeitsrangordnung zusammengefaßt; dort wird neben der zu prüfenden Hauptleistung eine gleichzeitige meßbare Nebenleistung verlangt; je mehr diese letztere beeinträchtigt wird, desto schwieriger ist die erstere.

Die ersten experimentellen Untersuchungen über die Schwierigkeit von einfachen Additionsaufgaben hat Jenson (12) an Kindern angestellt. Neben Beckmann (2), der seine Methode den untersuchten zwei- bis sechsjährigen Kindern angepaßt hatte, wurde die Rechenfertigkeit Jugendlicher noch in einer umfangreichen Untersuchung von Korn (13) geprüft. O. Lipmann (19) weist in seinem Sammelbericht auf die Arbeiten von Stown, Courtis und Brown hin. Im Gegensatz zu allen diesen Arbeiten hat Pauli (28) seine Untersuchungen mit Hilfe der Fehlerstatistik an Erwachsenen durchgeführt. Die Messung der Rechenreaktionszeiten verwendeten bereits Brandenburger (3) und Ranschburg (34, 35, 36).

In den bisherigen Untersuchungsergebnissen sind Widersprüche aufgetreten, d. h. bei den einzelnen Autoren weichen die Schwierigkeitsrangordnungen voneinander ab (S. 11, 17). Im Hinblick auf die Klärung dieser Frage hat die vorliegende Arbeit vor allem zwei Aufgaben zu erfüllen:

1. Das Hauptziel der Untersuchung ist die schon erwähnte Auffindung von objektiven allgemeingültigen Schwierigkeitsgraden oder noch enger umschrieben: Bis zu welcher Feinheit dürfen objektive Schwierigkeitsunterschiede aufgestellt werden, ohne daß sie durch individuelle Rangordnungen zum Teil wieder aufgehoben werden müssen?
2. Die beiden verwendeten Methoden, die Fehlerstatistik und die Zeitmessung sollen auf ihre Ergiebigkeit hin näher untersucht werden.

Die Lösung der gestellten Aufgaben soll durch eine Verfeinerung und einen Vergleich der beiden Methoden erreicht werden. Die Verfeinerung bezweckt die Schaffung möglichst günstiger Voraussetzungen für die Wahrscheinlichkeitsrechnung, nämlich strenge Gleichmäßigkeit aller Versuchsbedingungen und eine große Zahl von Vpn bzw. Einzelversuchen. Der fehlerstatistische Teil lehnt sich in seiner Anordnung an die Arbeit von Pauli (28) an. Die Verfeinerung gegenüber der dort angeführten Methode besteht

1. in der größeren Anzahl von Vpn: 57 Studenten bzw. Studentinnen rechneten über 150 000 Aufgaben, die ungefähr 4000 Falschlösungen einschließen;
2. in der Verwendung von genormten, eigens für die Zwecke der Statistik ausgearbeiteten Rechenbögen an Stelle der Kraepelinschen Rechenhefte (S. 7);

3. in der Auswertung der Verbesserungen; diese wurden hier in ihrer Verteilung auf die einzelnen Aufgaben und in ihrer Beziehung zu den reinen Fehlern zum ersten Mal eingehend untersucht.

Die Zeitmessung wurde mit Handstoppung und einfachem Tachistoskop durchgeführt (S. 8). Im Sinne des tachistoskopischen Zeitmeßverfahrens lag es auch, daß geräuschlos und soweit wie möglich unwissentlich gestoppt wurde. Auch hier wurde neben der strengen Gleichmäßigkeit der Vorbedingungen auf eine große Zahl von Einzelversuchen Wert gelegt: Mit 5 Vpn wurden über 13 000 Zeitmeßversuche durchgeführt. Außerdem wurde die Methode durch verschiedene Versuchsreihen ergänzt:

Um die Fehlerquellen, die durch das ungenaue Stoppen des Vl entstehen könnten, zu umgehen, wurde das Reaktionstachistoskop verwendet (S. 8).

Zur genaueren Bestimmung der Rechenzeiten wurden die Lesezeiten gesondert gemessen.

Außerdem wurden die Zwischenlesezeiten gestoppt (zwischen den zwei zu lesenden Zahlen).

Es muß auch zu den Fragen der individuellen Leistungsschwankung, der Übungsfähigkeit und der Gewöhnung, die im engsten Zusammenhang mit den hier zu erörternden Problemen stehen, Stellung genommen werden. Es ist kein Beweis gegen das Schwierigkeitsverhältnis einer Aufgabe zu einer anderen dadurch erbracht, daß die eine Aufgabe infolge der Leistungsschwankung bald schwieriger, bald leichter als die andere Aufgabe erscheint; beide Aufgaben müßten unter genau denselben Leistungsvoraussetzungen geprüft werden und es würden sich höchstwahrscheinlich Schwierigkeitsstufen feststellen lassen; diese Vorbedingung kann aber nie restlos erfüllt werden. Auf diese Fragen soll jedoch hier nicht näher eingegangen werden; es wird im Gegenteil versucht, den Einfluß der Leistungsschwankung durch möglichste Häufung der Versuche nach den Gesetzen der Wahrscheinlichkeitsrechnung abzuschwächen. Auch die durch die Übung bewirkte Schwierigkeitsverminderung ist nicht Gegenstand der Untersuchung. Werden die Aufgaben verschieden stark in ihrer Schwierigkeit vermindert, so erscheint das geradezu als ein Ausdruck ihrer verschiedenen Schwierigkeit, denn Erhöhung der Leistung ist nur bei relativ schwierigen Aufgaben wahrscheinlich und nötig. Die Schwierigkeitsstufen könnten unter Umständen durch extreme Übung im Einzelfall auf nicht mehr nachweisbare Unterschiede zusammengedrängt werden. Da bei Vp A (S. 22), die sehr geübt ist, noch große Belastungsunterschiede auftreten, ist eine völlige Nivellierung der Belastungen bei diesen einfachen Additionen nicht zu erwarten.

Die vorliegende Untersuchung entstand auf Veranlassung von Herrn Prof. Dr. Richard Pauli, dem ich für die vielen Anregungen und Ratschläge während der Arbeit hier herzlich danken möchte.

Durchführung der Versuche.

Versuchsmaterial.

Die Methode der Fehlerstatistik wurde mit Hilfe der von Pauli ausgearbeiteten genormten Rechenbögen durchgeführt; jeder Bogen besteht aus dickem, weißem Papier von 51,5 × 34 cm; jede Seite enthält die Möglichkeit zu 2000 fortlaufenden

Additionen: 40 Reihen zu je 50 Aufgaben. (Diese außerordentliche Stundenleistung wurde nur von 1 Vp erreicht.) Diese Rechenbögen wurden den von Kraepelin angegebenen Rechenheften vorgezogen, weil sie die Gleichartigkeit der Arbeitsvoraussetzung bei jeder Aufgabe besser gewährleisten: Die Schreibarbeit wird hier nicht mehr teilweise behindert (am inneren Rande des Heftblattes); während einer Stunde ist nur eine einmalige Wendung notwendig (bei den Kraepelinrechenheften mußte öfters die Seite gewechselt werden); der Hauptvorteil besteht aber in der gleichmäßigen Verteilung der Zahlen 2—9 und ihrer Zusammenstellungen; durch 0 oder 1 ist jeweils gekennzeichnet, wenn die Abfolge der möglichen Kombinationen durchlaufen ist. Es kann für die Auswertung der gerechneten Aufgaben angenommen werden, daß jede Aufgabe gleich oft gerechnet werden mußte.

Bei der Anwendung der Zeitmeßmethode wurden eine $^1/_{10}$ sec und eine $^1/_{100}$ sec Stoppuhr, eine $^1/_5$ sec Zweizeigerstoppuhr und die von Pauli angegebenen Apparate, ein einfaches Tachistoskop, ein Reaktionstachistoskop mit Kinnschlüssel und den dazugehörigen Lesekarten, eine $^1/_{100}$ sec Zweiuhrenstoppvorrichtung und eine Versuchswand benützt. Das Tachistoskop ist ein photographischer Momentverschluß mit verstellbarer Expositionszeit. (Nähere Beschreibung 29.) Das Reaktionstachistoskop in Verbindung mit dem Kinnschlüssel schaltet die Reaktionszeit des Vl aus. Mittels des Kinnschlüssels wird die Unterkieferbewegung, die mit jedem Sprechbeginn verknüpft ist, automatisch (durch Stromschluß) registriert. Das Kinn wird auf eine ihm angepaßte waagrechte Unterlage leicht aufgelegt; es kann dabei beliebig gegen eine dazu senkrechte Stütze gedrückt werden; beim Senken des Unterkiefers wird die waagrechte Unterlage niedergedrückt und die Uhr im Reaktionstachistoskop automatisch gestoppt. Die Zeitschwankungen, die in der Unterkieferbewegung bei verschiedenen Lauten auftreten, sind nicht größer als die für denselben Laut bei verschiedenen Reaktionen (24). Die 64 Additionsaufgaben (alle möglichen Zweizahlenkombinationen der Zahlen 2—9 in untereinanderstehenden Reihenfolgen) befinden sich auf 16 Lesekarten (8 × 4 cm), und zwar je zwei auf einer Seite. Die $^1/_{100}$ sec Zweiuhrenstoppvorrichtung ermöglicht es, 2 $^1/_{100}$ sec Stoppuhren zu gleicher Zeit geräuschlos in Gang zu bringen und einzeln abzustoppen *). Eine Versuchswand verdeckte den Vl völlig für die Vp.

Versuchsanordnung.

Die Versuche nach der Methode der Fehlerstatistik wurden mit 57 erwachsenen Vpn durchgeführt (39 Studenten und 18 Studentinnen). Es wurde dabei für gleichmäßige äußere Arbeitsbedingungen gesorgt (d. h. Stillzimmer und bequeme Schreib-

*) Zu beziehen durch die Firma M. Sendtner A.-G. München.

gelegenheit). Die Vpn hatten die Aufgabe, eine Stunde lang ununterbrochen an Hand der erwähnten Rechenbögen mit höchster Anspannung und größtmöglichster Rechengeschwindigkeit unter Verzicht auf jedes Nachrechnen und Verbessern zu addieren. Es durfte dabei auch nicht gesprochen oder vor sich hingemurmelt werden. Innerhalb der einzelnen senkrechten Zahlenreihen waren der Reihe nach von oben nach unten jeweils zwei aufeinanderfolgende Zahlen zu addieren, und ihr Ergebnis war rechts neben die Lücke der beiden Zahlen zu schreiben; bei Summen von 10 und darüber wurden die selbstverständlichen Zehnerzahlen ausgelassen; in jeder Reihe mußten die Einzeladditionen oben von neuem begonnen werden, es durfte also nicht die letzte Zahl einer Reihe zur ersten der folgenden gezählt werden. Da die Versuche hinsichtlich dieser Bedingungen ordnungsgemäß durchgeführt wurden, sind die auftretenden Fehler mit großer Wahrscheinlichkeit auf innere Beweggründe zurückzuführen.

Zu den Versuchen nach der Zeitmeßmethode hatten sich 5 erwachsene Vpn zur Verfügung gestellt (3 männliche und 2 weibliche), und zwar Frl. stud. phil. G. Burkhardt und A. Wittmann und die Herren Dr. W. Becker, stud. mus. F. Hafner und Lehrer K. Semmer, denen ich für die Mühe herzlich danken möchte. Alle 5 Vpn beteiligten sich an den Versuchen mit dem einfachen Tachistoskop und einer $^1/_{100}$ sec Stoppuhr; die Lesezeit und die Additionszeit (= Lese- und Rechenzeit) wurden dabei getrennt voneinander gemessen. Der Augenabstand war 25—30 cm (maximale Sehschärfe). Die Anordnung hieß: Bitte, die Zahlen von oben nach unten lesen und sofort aussprechen! (bzw. ... zusammenzählen und ihr Ergebnis sofort aussprechen). Auf ein Vorsignal (Bitte) wurde nach etwa 2—3 sec das Hauptsignal (Jetzt) gegeben und dann nach etwa ½—1 sec die Lesekarte exponiert (Expositionszeit $\sim {}^1/_{\tau}$ sec). Eine Versuchsreihe (alle 64 Kombinationen als Lese- oder Additionsaufgaben) nahm ungefähr 20—25 min in Anspruch, einschließlich 0—3 Pausen je nach dem Ermüdungszustand der Vp. Anfänglich hatte die Vp an einem Tage nur 1 Versuchsreihe zu lesen bzw. zu rechnen, später auch 2 und 3. Die Zeit zwischen Expositionsbeginn und zwischen Beginn des Aussprechens der ersten Zahl bei den Leseversuchen bzw. der Summe bei den Rechenversuchen stoppte der Vl ohne Wissen der Vp geräuschlos (in der Tasche). Ranschburg hatte jeweils das Ende des ausgesprochenen Ergebnisses festgestellt; dieser Zeitpunkt ist aber bedeutend schwieriger genau zu bestimmen als der Anfang; außerdem verlängert sich bei zweistelligen Summenergebnissen die Rechenzeit um die für den zweiten Teil des Wortes beanspruchte Zeit (fünf z e h n) und die verschieden lang ausgesprochenen Zahlwörter können die Rechenzeiten trüben.

Anschließend daran diente das Reaktionstachistoskop zur Bestimmung der Lese- bzw. Additionszeiten von 3 Vpn (2 männlichen und 1 weiblichen). Nach je 3—5 Vorversuchen erfolgte die

Durchführung in derselben Weise wie mit dem einfachen Tachistoskop.

Eine Ergänzung fanden die Leseversuche noch durch weitere Versuche mit 2 Vpn. Nach ergebnisloser Erprobung der Zweizeigerstoppuhr wurde mit Hilfe der Zweiuhrenstoppvorrichtung, die bei diesen Versuchen zum ersten Mal mit Erfolg Verwendung fand, geprüft, ob wesentliche Unterschiede in den Zwischenzeiten (zwischen Beginn des Aussprechens der ersten und Beginn des Aussprechens der zweiten Zahl) festzustellen wären. Im übrigen war auch hier die Anordnung dieselbe wie bei den beiden vorher angegebenen Versuchsgruppen.

Einzelauswertung der beiden Versuchsgruppen.

a) Auswertung der Fehlleistungsstatistik (1. Methode).

Die 57 Vpn haben zusammen 153 453 Additionen ausgeführt; dabei wurden 4007 Fehlleistungen (2060 Fehler und 1947 Verbesserungen) erzielt. Das ergibt eine durchschnittliche Belastung jeder Aufgabe mit 62 Fehlleistungen (32 Fehler und 30 Verbesserungen). Die Abbildung (S. 25) zeigt die Verteilung der Fehlleistungen auf die 64 Aufgaben in der Rangordnung ihrer Belastung. Als auffällig schwach belastete Aufgaben ergeben sich dabei 5 + 5 mit 10 Fehlleistungen, 3 + 3 mit 16, 2 + 2 mit 17, 5 + 2 mit 20, 7 + 7 mit 24 und 4 + 4 mit 26 Fehlleistungen, während 3 + 9 mit 116, 8 + 5 mit 119, 3 + 8 und 5 + 7 mit 145, 8 + 3 mit 170 und 7 + 5 mit 177 Fehlleistungen die Höchstbelastungen aufweisen. Die außerordentlich große Differenz zwischen schwächster und stärkster Belastung (167) schafft die Grundlage für die Brauchbarkeit dieser Methode. Die für den Ausbau der Methode wichtige Frage ist nun die: Bis zu welcher unteren Grenze sind Unterschiede in der Belastung noch von Bedeutung? Man könnte zu diesem Zweck etwa gleichgeartete Aufgaben, z. B. alle Aufgaben mit zwei gleichen Zahlen, also 2 + 2, 3 + 3 . . . . 8 + 8, 9 + 9 hinsichtlich ihrer Belastung untersuchen und die hierbei auftretenden Unterschiede für belanglos erklären (in unserer Reihe 24); jedes derartige Verfahren setzt aber bereits ein Urteil über die Schwierigkeit der Aufgaben voraus, das vorerst noch vermieden wird.

Eine Betrachtung der Übergänge von einer Aufgabe innerhalb der Reihe zur nächsten ergibt, daß diese relativ sehr klein sind; 12 mal bleibt der Sprung aus, während neben den 21 Einer-, 12 Zweier-, 4 Dreier-, 5 Vierer- und 4 Fünfersprüngen je 1 Sechser-, Siebener-, Zehner-, Fünfundzwanziger- und Sechsundzwanzigersprung festzustellen ist. Die größeren Sprünge liegen am Ende der Reihe und bedingen den exponentiellen Charakter der Rangordnungskurve.

Um das Verhältnis der eigentlichen Fehler zu den Verbesserungen zu klären, werden die Belastungen der Aufgaben mit Fehlern bzw. Verbesserungen verglichen. Geringe Fehlerbelastung weisen auf 5 + 5 mit 3 Fehlern, 2 + 2, 5 + 2, 3 + 3 mit je

6 Fehlern, 6 + 6 mit 8 und 8 + 8 und 4 + 3 mit je 10 Fehlern, hohe Fehlerbelastung 7 + 6 mit 61, 3 + 9 mit 67, 5 + 7 mit 72, 8 + 5 mit 80, 3 + 8 mit 93, 7 + 5 mit 100 und 8 + 3 mit 101 Fehlern.

Um einen Vergleich mit bisher gefundenen Versuchswerten der fehlerstatistischen Methode durchführen zu können, werden die Aufgaben, soweit sie auch in jenen Versuchen verwendet wurden, in einer Schw Rangordnung (siehe Fußnote Seite 13) zusammengestellt:

Eigene Versuche		Jenson	Korn	Pauli	Eigene Versuche		Jenson	Korn	Pauli
5 + 5	64	—	—	—	6 + 5	32	18	16	—
2 + 2	63	—	29	—	8 + 2	31	—	21	—
5 + 2	62	—	—	—	8 + 9	30	10	—	11
3 + 3	61	—	28	—	7 + 8	29	6	7	15
7 + 7	60	28	27	—	4 + 9	28	1	19	5
6 + 6	59	32	—	—	5 + 6	27	25	20	14
8 + 8	58	15	26	—	5 + 9	26	8	11	16
4 + 3	57	—	—	—	8 + 7	25	12	3	—
7 + 2	56	—	25	—	6 + 8	24	9	—	19
4 + 5	55	—	—	26	9 + 6	23	19	—	—
3 + 2	54	—	—	—	9 + 5	22	20	—	—
4 + 2	53	—	24	—	7 + 4	21	16	—	—
5 + 3	52	—	—	—	3 + 7	20	—	—	17
4 + 4	51	—	—	—	9 + 7	19	7	15	—
2 + 5	50	—	—	18	7 + 9	18	3	—	3
6 + 2	49	—	23	—	6 + 4	17	—	—	—
3 + 5	48	—	—	22	4 + 7	16	2	—	7
2 + 6	47	—	—	25	7 + 3	15	—	19	—
2 + 9	46	35	14	28	9 + 4	14	30	12	—
5 + 4	45	—	—	—	4 + 6	13	—	6	21
4 + 8	44	23	—	10	9 + 3	12	34	5	—
3 + 6	43	—	—	13	6 + 9	11	17	18	12
9 + 9	42	24	—	—	8 + 6	10	14	13	—
2 + 3	41	—	—	27	5 + 8	9	26	10	1
3 + 4	40	—	—	24	2 + 8	8	—	—	23
2 + 4	39	—	—	20	7 + 6	7	13	—	—
6 + 3	38	—	22	—	3 + 9	6	11	4	8
6 + 7	37	4	—	9	5 + 7	5	33	2	4
9 + 2	36	36	17	—	8 + 5	4	27	8	—
9 + 8	35	5	—	—	3 + 8	3	22	9	2
2 + 7	34	—	—	6	7 + 5	2	31	1	—
8 + 4	33	29	—	—	8 + 3	1	21	—	—

[Die Zahlen geben den Schw Rangplatz an (1 = schwierigste Aufgabe).]

Dabei ist zu beachten, daß die Werte von Pauli ohne Rücksicht auf die Reihenfolge der Summanden zusammengestellt sind (z. B. der Wert für 5 + 8 enthält die Werte von 5 + 8 und 8 + 5); außerdem, daß bei Korn die Aufgaben 19 mit 29 dieselbe Belastung haben (= 0).

Werden die Rangordnungen der einzelnen Verfasser mit der eigenen Rangordnung, die jeweils nur Aufgaben enthält, die auch in der damit verglichenen Ordnung vorkommen, verglichen, so ergeben sich folgende Korrelationskoeffizienten nach Spearman *): 0,12 (Jenson: eigene Versuche); 0,84

*) Im Rahmen der vorliegenden Untersuchung wurde zur Berechnung von Korrelationskoeffizienten nur die Formel von Spearman (47) benützt:

$$\varrho = 1 - \frac{6 \sum (u_1 - v_1)^2}{n\,(n^2 - 1)}$$

(Korn: eigene Versuche); 0,63 (Pauli: eigene Versuche). Der Vergleich ist nicht genau durchzuführen, da bei den einzelnen Autoren die verwendeten Aufgaben nicht ganz gleich sind. Jensons Werte stammen von je 45 Schülern, die in 6 aufeinanderfolgenden Jahren von ihm als Rechenlehrer geprüft wurden; er hat dabei schwierige Aufgaben, unter die er selbst z. B. auch 7 + 5 zählt (siehe Zusammenstellung!), besonders fleißig geübt. Es kam ihm nicht darauf an, objektive Schw Grade festzustellen: „Von objektiver Schw und Leichtigkeit gewisser Aufgaben kann also nicht die Rede sein" (12, S. 701). Korn (13) standen aus seinen Versuchen mit Kindern 284 F zur Verfügung, die sich auf 64 Aufgaben verteilen. Die Rangordnung von Pauli (28) ist das Ergebnis von Versuchen mit 11 Studenten bzw. Studentinnen. Die Werte von Beckmann (2), der 2—6jährige Kinder untersuchte, scheiden bei diesem Vergleich aus, da sich seine Summen nur zwischen 2 und 6 bewegten. Ebenso kann die in der Arbeit von Wirth erschienene Rangordnung nicht mitverglichen werden (46), da ihr als Schw Kriterium die „Fehlerdauer" (Gemeinsame Rangordnung aus den Werten der fehlerstatistischen und der Zeitmeßmethode) zugrundeliegt; außerdem ist für diese dort noch zum Vergleich mitangeführte Schw Rangordnung der Additionsaufgaben nicht ersichtlich, welches das zugrundegelegte Material ist. Die im Rahmen vorliegender Untersuchung sich ergebenden Gesetzmäßigkeiten stimmen aber gut mit den Ergebnissen von Wirth überein (s. u.); die beiden Arbeiten sind in der Anlage ähnlich, insoferne die Arbeit Wirths, die gleichzeitig mit vorliegender im Psychologischen Institut der Universität München entstanden ist, die Schw Verhältnisse an einfachsten Subtraktionen mit Hilfe der F Statistik und der Zeitmessung prüft.

Die Aufgaben mit den kleinsten Zahlen von Verbesserungen sind 5 + 5 mit 7, 3 + 3 mit 10, 2 + 2 mit 11, 8 + 2, 3 + 4 und 4 + 4 mit 12 und 3 + 5 mit 13 Verbesserungen, mit den höchsten Zahlen von Verbesserungen 3 + 9 und 6 + 9 mit 49, 3 + 8 mit 52, 9 + 3 mit 57, 8 + 3 mit 69, 5 + 7 mit 73 und 7 + 5 mit 77 Verbesserungen.

Nach den kleinsten und größten Werten in den Rangordnungen der F und V gemessen, möchte es scheinen, als sei kein grundsätzlicher Unterschied bei einem Vergleich der beiden Reihen festzustellen: Die drei kleinsten Werte sind jeweils 5 + 5, 2 + 2 und 3 + 3 und die zwei größten Werte 5 + 7 und 3 + 8. Da aber im Innern der Reihen Verschiebungen zu finden sind, wird der Übereinstimmungsgrad der beiden Reihen genau festgestellt: Auf Grund der Höhe des Korrelationskoeffizienten (0,74) erscheint es gerechtfertigter, die zwei Reihen gemeinsam als getrennt voneinander zu behandeln. (K.k. für F und Fehlleistungen sogar 0,95). Damit werden bessere Voraussetzungen für die Wahrscheinlichkeitsrechnung (d. h. eine größere Anzahl von Falschlösungen auf die einzelnen Aufgaben) erreicht. Die Fehlleistungskurve (S. 25) wird also als grundlegende Rangordnungskurve der fehlerstatistischen Methode (1. Methode) betrachtet. Trotzdem soll das Verhältnis der beiden Reihen bei jeder sich bietenden Gelegenheit näher betrachtet werden, um das Verhältnis der F zu den V genauer zu erklären: Die Spanne für F beträgt 32, d. h. die kleinste F zahl (5 + 5) ist in der Differenz zwischen größter und kleinster F zahl ungefähr 32mal enthalten, während die Spanne für V ungefähr 10 beträgt, also ungefähr nur den 3. Teil der Reichweite der F, d. h. es sind bei den F deutlichere

Unterschiede vorhanden. Zum gleichen Ergebnis kann man auch auf anderem Wege gelangen: Bei einer Nachprüfung, welche Aufgaben mehr F als V verursachten, ergibt sich, daß die am stärksten belasteten Aufgaben größtenteils mehr F als V, während die am schwächsten belasteten Aufgaben — obwohl die Gesamt V zahl kleiner ist als die Gesamt F zahl — größtenteils mehr V auf sich vereinigen: Die F sind also zur Erfassung der Schwierigkeit*) geeigneter als die V, d. h. sie sind mehr durch die den Aufgaben innewohnende Schwierigkeit verursacht worden als die V, die wohl in erhöhtem Maße Leistungsschwankungen und innere und äußere Ablenkungen während der Versuchsstunde zum Ausdruck bringen; allerdings frägt sich, ob die größere Anzahl von V nicht auch eine stärker wirksamwerdende Selbstbeobachtung und Selbstkritik anzeigen kann.

Die Voraussetzung für die Stichhaltigkeit dieser Feststellung über das Verhältnis der F und V ist der hohe Übereinstimmungsgrad ihrer Rangordnungen. Eine merkwürdige Ausnahme machen dabei die Aufgaben $2+8$, $4+6$ und $3+7$; da ihre Rangplätze in der F Statistik zu den Extremwerten zählen und dabei beträchtlich höher liegen als in der V Statistik, kann die erhebliche Überzahl der F gegenüber den V bei diesen Aufgaben nicht zur Stützung des Satzes herangezogen werden, daß F zur Erfassung der Schwierigkeit geeigneter sind als V; Aufgaben, die in ihrer Summe 10 ergeben, scheinen also dabei eine Ausnahme zu machen.

Im Anschluß daran soll untersucht werden, wie weit sich Unterschiede bei einer Auswertung nach Geschlechtern geltend machen. Die weiblichen Vpn haben im Durchschnitt mehr Aufgaben gerechnet als die männlichen (2804 : 2641) und dabei auch eine geringere durchschnittliche Fehlleistungszahl erreicht als die männlichen (68 : 71); es ist also zunächst eine etwas größere Leistungsfähigkeit der weiblichen Vpn festzustellen. Es wird nun die Verteilung der F und V auf die beiden Geschlechter verglichen. Die beiden zu vergleichenden Gruppen enthalten jeweils dieselbe Zahl von F bzw. V. Es zeigt sich dabei eine auffällige Änderung der Schw Stufen. Um die Frage besser zu klären wird ein weiteres umfassenderes Schw Kriterium verwendet, d. h. es wird nachgeprüft, ob die Schw Steigerung bei Aufgaben mit einstelligen Summen zu denen mit zweistelligen Summen bei beiden Geschlechtern wesentlich verschieden ist. Es ergibt sich dabei eine Belastung mit 91 F für Aufgaben mit einstelligen Summen gegenüber 424 F bei Aufgaben mit zweistelligen Summen für männliche Vpn (149 : 534 V); bei den weiblichen Vpn, die ungefähr dieselben Rangplätze wie die männlichen Vpn in den Rangordnungen der Vpn nach ihrer Belastung mit F bzw. V einnehmen, ist eine Belastung von 67 F (151 V) für Aufgaben mit einstelligen Summen gegenüber 448 F (bzw. 532 V) bei Auf-

*) Der Einfachheit halber wird von jetzt ab bereits von „Schwierigkeit" gesprochen, d. h. höhere oder geringere Belastung mit F (bzw. V) wird gleichgesetzt mit größerer oder geringerer Schw. Auf Seite 12 wird unabhängig von diesen und folgenden Feststellungen nachgewiesen, daß die in der 1. und 2. Methode verwendeten Kriterien zur Erfassung der Schw stichhaltig sind.

gaben mit zweistelligen Summen festzustellen. Es ergibt sich also für die Verhältnisse bei der F Statistik ein deutliches Überwiegen des Schw Anstieges für weibliche Vpn (40 %). Die Verteilung der V zeigt ein geringes Übergewicht des Schw Anstieges bei männlichen Vpn an (3%). Da aber die Methode der F Verteilung als ergiebiger zur Bestimmung von Schw Verhältnissen erkannt wurde (S. 13), wird ihr eindeutiges Ergebnis höher eingeschätzt. Allerdings erscheint das große Übergewicht vor allem im Hinblick auf die zunächst festgestellte geringe Überlegenheit der weiblichen Leistungsfähigkeit als etwas hoch; um nun besser entscheiden zu können, ob es sich um Schwankungen, die sich durch die Betrachtungsweise ergeben, oder um eine grundsätzliche Erscheinung handelt, soll der Schw Anstieg der beiden Gruppen näher untersucht werden; es wird bei jeder Gruppe (Aufgaben mit einstelligen und zweistelligen Summen) berücksichtigt, aus wievielen Einzelaufgaben sich das Ergebnis zusammensetzt. Die Differenz aus zwei- und einstelligen Summen für männliche Vpn beträgt 76,8 F — 23,8 F = 53 F (bzw. 106,6 V — 36,2 V = 70,4 V), die für weibliche Vpn 85,5 F — 17,3 F = 68,2 F (bzw. 102,5 V — 36,8 V = 65,7 V). Es ergibt sich hier also ungefähr dasselbe wie bei der vorhergehenden Betrachtung; das Übergewicht der weiblichen Vpn würde hier zwar noch etwas höher liegen. Da aber der Schw Anstieg bei beiden Gruppen sehr viel höher liegt als der für die Gesamtfehlerzahl errechnete, wird der Grund für den noch größeren Schw Anstieg beim weiblichen Geschlecht wohl in dem ungleichmäßigen Schwanken der männlichen und weiblichen Gruppe um den als Durchschnitt gefundenen Wert zu suchen sein; denn bei Annahme dieses Übergewichts beim weiblichen Geschlecht müßte das auch in der Rangordnung der Unterschiede derselben Aufgaben bei beiden Gruppen deutlicher zum Ausdruck kommen, d. h. es müßten die in der Statistik am stärksten belasteten Aufgaben eindeutig auch die größeren Unterschiede ergeben; das ist aber nicht der Fall. Es zeigt sich grundsätzlich aber die geringere Leistungsfähigkeit beim weiblichen Geschlecht. Das zu Beginn dieses Abschnittes erwähnte Übergewicht der weiblichen Vpn hinsichtlich der durchschnittlichen Aufgabenzahl und der durchschnittlichen Belastung mit Fehlleistungen kann nicht zum Beweis dagegen angeführt werden; denn der Unterschied ist so gering, daß er als eine durch die Methode bedingte Schwankung aufgefaßt werden kann.

Bei der Rangordnung der F sowie der der V nach ihrer Verteilung auf die einzelnen Vpn zeigt sich, daß die Extremwerte nach oben und unten auf der männlichen Seite zu suchen sind.

Brandenburger (3) fand, daß die Knaben wohl durchschnittlich kleinere Reaktionszeiten haben, daß aber ihre Leistungen eine geringere Einheitlichkeit aufweisen als die der Mädchen. Pauli stellte fest (31, S. 153): „Die Streuung (die interindividuelle Variation) ist beim männlichen Geschlecht größer als beim weiblichen."

Auffällig ist die Verteilung der F und V auf die beiden Geschlechter. Unter der Gruppe, die mehr F als V auf sich vereinigt, nehmen die weiblichen Vpn mit 21% Anteil, während sie in der Gruppe, die mehr V als F aufweist, mit 43 % vertreten sind. Die Frau neigt also bedeutend mehr dazu V zu machen, während diese beim Mann zugunsten der F stark zurückgedrängt sind. Wieweit das ein Ausdruck dafür ist, daß weibliche Vpn subjektiven Leistungsschwankungen innerhalb der Versuchsstunde und inneren und äußeren Ablenkungen mehr unterworfen sind, oder auch dafür, daß die unbewußte Selbstkontrolle und -kritik bei ihnen stärker wirksam ist als bei männlichen Vpn kann nicht eindeutig festgestellt werden.

Mit der Betrachtung der verschiedenen Leistungsfähigkeit bei beiden Geschlechtern werden Zusammenhänge, die zwischen Rangplatz in der Rangordnungskurve (die Belastung aller Vpn mit F bzw. V) und Leistungsfähigkeit bestehen, aufgedeckt: Die Höhe des Rangplatzes bringt scheinbar nicht immer auch die Abhängigkeit von der Schw klar zum Ausdruck, z. B. 5 + 7 und 7 + 5 mit zusammen 116 F haben 20 F mehr bei männlichen Vpn, 4 + 5 und 5 + 4 mit zusammen 12 F 6 F mehr bei weiblichen Vpn, dagegen 3 + 6 und 6 + 3 mit zusammen 20 F 12 F mehr bei männlichen Vpn und 5 + 8 und 8 + 5 mit zusammen 89 F 11 F mehr als bei weiblichen Vpn (ähnlich bei V). Eine Begründung dafür kann man in der Verschiedenheit der männlichen und weiblichen Rangordnung erblicken, abgesehen von der durch die Aufspaltung des statistischen Materials sich ergebenden größeren Schwankungsbreite. Jedenfalls taucht dabei aber auch die allgemeine Frage auf, ob die höhere oder geringere Belastung mit F bzw. V immer eine größere oder geringere Schw Abhängigkeit ausdrücken muß. Um diesen Sachverhalt zu klären, werden die Vpn mit den höchsten F bzw. V Zahlen den Vpn mit den niedrigsten F bzw. V Zahlen gegenübergestellt. Jede Gruppe enthält die gleiche Anzahl von F bzw. V: Die erste Gruppe erfährt eine Steigerung in der Belastung von 1 : 5,6 (1 : 3,3 bei V), wenn das Verhältnis der auf einstellige zu den auf zweistellige Summen entfallenden F Zahlen gebildet wird, während in der zweiten Gruppe die Steigerung 1 : 3,9 (1 : 3,2 bei V) eintritt. Vpn mit einer hohen Belastung von F bzw. V sind also mehr von den Faktoren abhängig, die mit der Schw in Beziehung stehen als Vpn mit niederen F bzw. V Zahlen oder: Die höhere Belastung mit F bzw. V zeigt zugleich die größere Abhängigkeit von Schw Verhältnissen an. (Siehe S. 26). Damit ist ein Beweis für die Stichhaltigkeit der F bzw. V statistischen Methode erbracht Dieser Zusammenhang gilt anscheinend nicht für weibliche Vpn (s. o); es wird also noch nachgeprüft, ob für das weibliche Geschlecht dieser Zusammenhang auch besteht: Es werden der weiblichen Vp mit der höchsten F (bzw. V) Belastung einige weibliche Vpn mit geringer F (bzw. V) Belastung gegenübergestellt,

die aber zusammen dieselbe Zahl von F bzw. V aufweisen wie die Vp mit der höchsten Belastung. Es ergibt sich auch für das weibliche Geschlecht die Abhängigkeit von der Schw entsprechend der absoluten Höhe der Belastung mit F bzw. V.

Gegen das Ergebnis der Auswertung nach verschiedenen Leistungstypen kann eingewendet werden, daß die Extremwerte nur teilweise vorausgesetzt sind; um bei beiden Gruppen die gleichen Zahlen von F und V zu gewinnen, wurden 6 Vpn aus dem Ende der F bzw. V Rangordnung 29 Vpn aus dem Anfang der jeweiligen Rangordnung gegenübergestellt. Es sind also die Extremwerte nur in der Gruppe der hohen F bzw. V Zahlen erfaßt worden. In einer zweiten Auswertung werden die Gruppen aus je 28 Vpn aus den ersten Hälften der Rangordnungen (F und V) und aus je 28 Vpn aus ihren zweiten Hälften zusammengesetzt. Die Auswertung wurde wie oben (S. 15) vorgenommen. Das Ergebnis beider Zusammenstellungen bestätigt die oben angeführte Behauptung: Die Steigerung in der Belastung der ersten Gruppe (hohe Belastung) beträgt 1 : 7,4, während sie für die zweite Gruppe 1 : 4,9 beträgt. Die Erhöhung dieses Anstieges wird gegenüber dem entsprechenden der obigen Auswertung (1 : 5,6 und 1 : 3,9) durch die ungleiche Verteilung der F Zahlen auf beide Gruppen bedingt. Entsprechend liegen die Verhältnisse bei den V: Anstieg in der ersten Gruppe 1 : 4,2, in der zweiten Gruppe 1 : 4,5 (s. o. 1 : 3,3 und 1 : 3,2). Die Tatsache, daß das Verhältnis für V hier ebenso das umgekehrte ist wie bei der Untersuchung der Geschlechtsunterschiede (S. 14) weist darauf hin, daß dort das umgekehrte Verhältnis nicht auf den Zusammenhang der V mit dem weiblichen Geschlecht zurückzuführen ist.

Ein Vergleich der absoluten Aufgabenzahlen der einzelnen Vpn mit den dazugehörigen Fehlleistungszahlen läßt keine Ordnung etwa in dem Sinne erkennen, daß Vpn mit hoher Aufgabenzahl auch eine entsprechend hohe Anzahl von Fehlleistungen aufzuweisen hätten (z. B. eine Vp mit 3427 Aufgaben hatte 147 Fehlleistungen, während eine andere Vp mit 1950 Aufgaben 27 Fehlleistungen aufwies; dagegen hatte eine Vp mit 3407 Aufgaben 25 Fehlleistungen, während eine andere Vp mit 1679 Aufgaben 93 Fehlleistungen auf sich vereinigte). Es soll nun genauer untersucht werden, ob die Vermutung, die Zahl der Aufgaben stehe nicht in direktem Zusammenhang mit der Zahl der Fehlleistungen, richtig ist. Werden 6 Vpn aus der ersten Hälfte der Rangordnung der absoluten Aufgabenzahlen 6 Vpn aus ihrer zweiten Hälfte gegenübergestellt (mit jeweils 2 weiblichen Vpn und 425 Fehlleistungen in jeder Gruppe), so ergibt sich eine Steigerung von 95 : 330 für hohe und 80 : 345 für niedere Aufgabenbelastung, wenn das Verhältnis der auf einstellige zu den auf zweistellige Summen entfallenden Fehlleistungszahlen gebildet wird. In einer weiteren Auswertung werden 28 Vpn der ersten Hälfte den 28 Vpn der zweiten Hälfte der Rangordnung der absoluten Aufgaben-

zahlen gegenübergestellt (Vp 29 wird dabei ausgeschieden). Dabei ergibt sich für die erste Hälfte (Vpn mit hoher Aufgabenzahl) ein Schw Anstieg von 1 : 4,2, für die zweite Hälfte ein solcher von 1 : 4,7. Da das Übergewicht zu gering ist, kann die oben angeführte Vermutung nicht eindeutig widerlegt werden. Im Hinblick auf den deutlichen Schw Anstieg bei der Auswertung nach Geschlechtern bzw. nach Gruppen mit hoher und geringer Belastung mit F bzw. V ist die Frage zu klären, ob der geringe Anstieg hier nur auf die veränderten Verhältnisse durch die Verwendung der Fehlleistungszahlen (dort F Zahlen) zurückzuführen ist. Bei der Verwendung der F Zahlen kommt nur ein geringes Übergewicht des Schw Anstieges bei den Vpn mit hoher Aufgabenzahl (1 : 5,8) gegenüber dem bei Vpn mit geringer Aufgabenzahl (1 : 5,3) zustande. Es kommt also zum Ausdruck, daß keine Abhängigkeit der Schw von der absoluten Aufgabenzahl festgestellt werden kann.

Es soll nun noch die Frage untersucht werden, welcher Art die F und V sind. Die folgende Übersicht gibt Zahl und Art der Abweichungen vom jeweils richtigen Summenergebnis nach oben (+) und unten (—) an:

Art und Zahl der Abweichungen	+7	+6	+5	+4	+3	+2	+1	−1	−2	−3	−4	−5	−6	−7
Fehler	5	7	8	13	74	345	640	368	286	86	70	33	44	26
Verbesserungen	8	9	4	26	76	242	415	422	271	50	62	41	25	8

Es ergaben sich in beiden Fällen Gauß'sche Kurven (mit 3 unbedeutenden Abweichungen), d. h. die Zahl der Abweichungen nimmt mit ihrer Größe ab. Hervorstechende Werte sind dabei + (—) 2 und + (—) 1, die zusammen über 4mal sowiel Abweichungen auf sich vereinigten als alle übrigen zusammen. Bei einer Gegenüberstellung der Verteilungen der stark- und schwachbelasteten Aufgaben (mit Fehlleistungen) zeigt sich, daß bei starkbelasteten Aufgaben die Abweichung um + 1 und — 1 bedeutend öfters auftritt als bei schwachbelasteten Aufgaben (900 : 434 bei + 1 und 634 : 412 bei — 1); im Gegensatz dazu ist von + (—) 2 bis + (—) 7 ein Überwiegen der Abweichungen schwachbelasteter über starkbelastete Aufgaben festzustellen; dieses Überwiegen ist außerdem auf der + Seite viel ausgeprägter als auf der — Seite. Diese Belastungsunterschiede zwischen + (—) 1 und den übrigen Werten bringen zum Ausdruck, daß die kleinen Abweichungen (+, — 1) ein mehr durch die Schw der Aufgabe bedingtes Versagen bedeuten als die größeren Abweichungen, die mehr durch andere Gründe verursacht sind. Bei der Einzelauswertung für F und V

ergibt sich grundsätzlich dasselbe Ergebnis; das Übergewicht der +, — 1 Werte ist dabei für F viel ausgeprägter als für V (860 : 150 bei F gegenüber 676 : 354 bei V); das entspricht der oben angeführten Behauptung, daß F zur Erkenntnis der Schw besser geeignet sind als V.

Um auf die Entstehungsarten der F bzw. V näher eingehen zu können, wird versucht, die einzelnen Abweichungen in typische Gruppen von Falschlösungen einzuteilen. Es kann dabei oft schwer oder überhaupt nicht entschieden werden — auch schon wegen der nur einstelligen Schreibweise der zweistelligen Summenergebnisse — in welche Gruppe die einzelne Abweichung eindeutig einzureihen ist; daher ist auch eine genauere quantitative Auswertung der Gruppen nicht möglich.

1. a) Lese- und Schreibfehler: Z. B. 3 + **3** (**5**) = 8; 7 + **3** (**5**) = 12; 9 + **3** (**2**) = 11.
   b) Klangbildverwechslungen: Z. B. **2** (**3**) + 3 = 6; **2** (**3**) + 5 = 8; **4** (**5**) + 4 = 9.
2. Zeitweiliges Aussetzen des Aufgabenbewußtseins:
   a) Perseveration: **7** + **7** = **7**; 2 + **8** = **8**; **3** + 2 = **3**; 9 + **3** = **13**; **3** + 9 = **13**.
   Dazu sind die Fälle zu zählen, bei denen das optische oder akustische Zahlbild des vorangehenden Summenergebnisses oder das der vorausgehenden oder der folgenden Zahl wirksam wird.
   b) Einfaches Weiterzählen: 2 + **8** = **9**; **4** + 2 = **5**; 3 + **4** = **5**.
   Weiterzählen in Intervallen: **4** + **6** = **8**.
   Einfaches Abwärtszählen: **6** + **5** = **4**; **7** + **6** = **5**;
   Ein Abwärtszählen entsprechend dem Aufwärtszählen etwa in folgender Form **7** + **3** = **6** tritt nicht auf, da die Tendenz des Abwärtszählens erst durch die beiden vorangehenden Zahlen ausgelöst werden muß: Das Aufwärtszählen zeigt sich also hier als das Ursprünglichere (vgl. starkes Übergewicht der + 1 Werte S. 17).
   Abwärtszählen in Intervallen: **9** + **6** = **3**; **9** + **7** = **5**; **7** + **5** = **3**; **6** + **4** = **2**.
   c) Verwechseln der Funktion: Abziehen bzw. Zerlegen: **7** + **2** = **5**; **2** + **7** = **5**.
   Vervielfachen und Teilen: **2** + **4** = **8**; **9** + **3** = **3**.
   Dazu noch Überspringen des Ergebnisses: **7** + **2** (= **5** — **2**) = **3**.
   Entsprechend bei richtiger Funktion: **2** + **2** (= **4** + **2**) = **6**; **2** + **3** = **5** + **3**) = **8**.

Die F sind größtenteils durch ein Nachlassen der Bewußtseinskontrolle bedingt; je größer nun die zu überwindende Schw ist, desto eher wird eine bereits vorhandene Disposition zum F führen. Die auftretenden Aufmerksamkeitsschwankungen können verschieden stark sein: Eine Vp weiß im Augenblick überhaupt nicht mehr, daß sie eine mathematische Funktion auszuführen hat und wiederholt eine vorhergehende Zahl oder zählt weiter; eine andere Vp ist sich noch bewußt, daß sie eine lineare mathematische Bewegung zu vollziehen hat (+ oder —), aber ohne Richtungsbewußtsein. Einer Vp ist geläufig, daß beim Dazuzählen einer einstelligen Zahl zu 9 über 10 hinaus die dazugezählte Zahl um 1 vermindert und der Zehner vorgesetzt werden muß, z. B. 9 + 4 = 13. Diese Verminderung um 1 kann sich in einer Ermüdungsschwankung des Bewußtseins in eine Vermehrung um 1 verwandeln, z. B. 9 + 4 = 15 oder 5 + 9 = 16. Die verschiedenen F Arten zeigen, daß die Aufgabenlösung auch bei fortschreitender Mechanisierung ein komplexer Vorgang ist, *der bis zu einem gewissen Grad immer eine tatsächliche intellektuelle Leistung darstellt*. Auch Korn (13) und Ranschburg (36) kommen zu dem Ergebnis, daß es kein mechanisches Rechnen, auch nicht bei den elementarsten Rechenarten gibt.

b) Auswertung der Zeitmessung (2. Methode).

Der Auswertung nach der 2. Methode liegen außer den Vorversuchen 13 056 einzelne Zeitmessungen, und zwar 5952 Messungen der Lesezeit (n = 93 für eine Aufgabe) und 7104 Messungen der Rechenzeit (n = 111 für eine Aufgabe) zugrunde. Als Mittelwert aus den einzelnen Versuchen wurde jeweils das Stellungsmittel genommen, um Extremwerte auszuschalten. Um die „absolute Rechenzeit" einer Rechenaufgabe feststellen zu können, mußte die jeweils für eine Aufgabe beanspruchte Lesezeit von der gesamten Rechenzeit in Abzug gebracht werden, d. h. es mußte untersucht werden, ob eine auftretende Rechen Schw nicht aus einer Lese Schw herzuleiten ist. Absolute Rechenzeit ist dabei nur in dem Sinne von Rechenzeit ohne die dazugehörige Lesezeit zu verstehen. In ihr sind noch andere Zeiten unberücksichtigt geblieben: Wenn der Zeitpunkt für das Stoppen der Uhr und die Auslösung des tachistoskopischen Verschlusses im Idealfall als gleichzeitig angenommen wird, ist neben der Lesezeit noch die von der Vp benötigte Zeit zum Bilden und Aussprechen der Zahl miteingeschlossen; beim Stoppen mit der Hand kommt außerdem noch die vom Vl benötigte Zeit zum Auffassen des Hörbildes sowie die für das Stoppen der Uhr benötigte Zeit hinzu. Es kann damit auch nicht die Frage geklärt werden, ob das Lesen ohne die Aufgabenstellung der Addition nicht von sich aus in kürzerer Zeit zu bewältigen sein wird als das Lesen im Hinblick auf eine zu lösende Aufgabe. Es würde über den Rahmen der vorliegenden Untersuchung hinausgehen, eine genaue, im Sinne einer Erfassung der absoluten Rechenzeiten notwendige Scheidung zwischen eigentlichem Leseakt und einer vertieften oder zum mindesten andersartigen Erfassung der Zahl, die durch die zu lösende Aufgabe notwendig wird, zu finden, etwa in Form von Zahlreproduktionsversuchen. Für unsere gröberen Leseversuche kam es darauf an, auffällige Zeitunterschiede beim Lesen der einzelnen Aufgaben festzustellen, die sich dann notwendig auch beim Addieren geltend machen müssen. Es ergibt sich dabei, daß im großen und ganzen die Schwankungen der Lesezeiten, die durch die Methode verursacht sind, größer sind als die durch die Lese Schw verursachten Belastungen; da die durch die Methode bedingten Schwankungen aber sehr gering sind (s. u.), sind auftretende Unterschiede in den Lesezeiten für diese Versuche belanglos; wenngleich einzelne Zahlenzusammenstellungen mit längerer Lesezeit eine lange Rechenzeit beanspruchen, z. B. 6 + 9 bzw. 9 + 6; es ist aber nicht angängig, die absoluten Rechenzeiten so zu bilden, daß von der Rechenzeit jeder Aufgabe die für diese Aufgabe gefundene Lesezeit abzuziehen ist, da nicht entschieden ist, ob eine besonders hohe Rechenzeit immer durch eine hohe Lesezeit gekennzeichnet sein muß. Die Lesedauer (St M für die 3 Vpn zu-

sammen, n = 30) beträgt für die Versuche mit dem Reaktionstachistoskop 0,56 sec (A M = 0,56 sec); die Streuung ist sehr gering, und zwar m V ± 0,15 sec, ± 2,7 %. Für die 5 Vpn, die an den Versuchen mit einfacher tachistoskopischer Zeitmessung beteiligt waren, ergeben sich für das Lesen zweier Zahlen folgende Zeitwerte: St M 0,80 sec (A M 0,81 sec), m V ± 0,017 sec, ± 2,1 % (n = 28).

Diese Zahlen, die in der Arbeit von Wirth (46) bereits veröffentlicht sind, finden eine gute Bestätigung durch die dort gefundenen Werte; für die reaktionstachistoskopischen Versuche ergeben sich dort: St M 0,51 sec (A M 0,51 sec), m V ± 0,04 sec, ± 1,6 % (1 Vp, n = 3); für die einfachen tachistoskopischen Versuche: St M 0,74 sec, m V ± 0,02 sec, ± 2,7 % (6 Vpn, n = 13).

Diese geringe Streuung rechtfertigt die Aufstellung eines Gesamtstellungsmittels der Lesezeiten aller Aufgaben für jede Vp. Die absoluten Rechenzeiten unterscheiden sich also von den gestoppten nicht wesentlich, nur ihre Spanne ist größer.

Gegen diese Art der Erfassung der absoluten Rechenzeiten können insoferne Bedenken erhoben werden, als bei der Zeitstoppung immer nur die Zeit bis zum Aussprechen der ersten der beiden Zahlen gemessen wurde; es konnte nichts über die Schw, die beim Zusammenlesen der beiden Zahlen auftreten könnte, ausgesagt werden. Zu diesem Zweck wurden Zusatzleseversuche bei 2 Vpn mit Hilfe der Zweizeigerstoppuhr durchgeführt; da die Zwischenlesezeiten aber sehr gering waren, konnten sie nicht mit der nötigen Genauigkeit festgestellt werden. Es wurde daher die erwähnte Zweiuhrenstoppvorrichtung verwendet, die sich als sehr praktisch im Gebrauch erwies. Die Ergebnisse stützen die Aufstellung des Gesamtstellungsmittels; es zeigt sich bei der Auswertung nicht, daß Aufgaben mit langen Rechenzeiten auch große Zwischenlesezeiten aufzuweisen hätten.

Die Abbildung (S. 25) zeigt die Verteilung der absoluten Rechenzeiten auf die 64 Aufgaben in der Rangordnung ihrer Belastung. Als schwachbelastete Aufgaben ergaben sich dabei 5 + 5 mit $^{12}/_{100}$ sec, 4 + 4 mit $^{13}/_{100}$ sec, 2 + 2 mit $^{14}/_{100}$ sec, 7 + 7 und 3 + 3 mit $^{15}/_{100}$ sec und 8 + 2 mit $^{16}/_{100}$ sec, während 7 + 9 $^{39}/_{100}$ sec, 8 + 7 und 6 + 8 $^{40}/_{100}$ sec, 4 + 7 $^{41}/_{100}$ sec, 7 + 8 $^{42}/_{100}$ sec und 6 + 9 $^{48}/_{100}$ sec beanspruchten (die Zeiten sind aufgerundete Durchschnittszeiten für Hand- und Reaktionstachistoskopische Zeitmessung). Die Differenz zwischen schwächster und stärkster Belastung beträgt $^{36}/_{100}$ sec (Spanne: 3), also groß genug für eine stichhaltige Auswertung.

Eine Betrachtung der Übergänge von einer Aufgabe innerhalb der Reihe zur nächsten ergibt, daß diese äußerst gering sind: 37 mal bleibt der Sprung aus, während neben den 22 Einer- und den 2 Zweier-Sprüngen je 1 Vierer- und 1 Sechser-Sprung festzustellen ist.

Der Durchschnittswert der absoluten Rechenzeiten beträgt für 1 Vp und 1 Aufgabe bei den handgestoppten Versuchen 0,26 sec und für die reaktionstachistoskopischen Versuche auch 0,26 sec;

die durchschnittlichen Lesezeiten (und damit auch die gesamten Rechenzeiten = Rechenzeiten + Lesezeiten) fallen von 0,80 sec (St M) bei den einfachen tachistoskopischen Versuchen auf 0,56 sec bei den reaktionstachistoskopischen Versuchen. Wirth (46) fand als Durchschnittswert für seine Schw Bestimmungen an einfachen Subtraktionen 0,93 sec; die Erhöhung der Rechenzeit bei Subtraktionen gegenüber Additionen stellte außer Korn (13) bereits Ranschburg fest (36): Als Durchschnittsbelastung seiner Versuche an Schulkindern fand er für Additionen 3,0 sec und für Subtraktionen 4,4 sec.

Einen Vergleich der eigenen mit der Rangordnung von Ranschburg (36) zeigt folgende Zusammenstellung (es wurden die in beiden Versuchen gemeinsam verwendeten Aufgaben gegenübergestellt):

Reihenfolge in abst. Schw	Ranschburg	Eigene Versuche	Reihenfolge in abst. Schw	Ranschburg	Eigene Versuche
1	6 + 5	8 + 3	13	4 + 3	3 + 4
2	6 + 4	3 + 5	14	2 + 7	4 + 5
3	4 + 5	2 + 7	15	7 + 2	5 + 4
4	5 + 4	6 + 5	16	4 + 2	4 + 2
5	3 + 4	7 + 2	17	6 + 2	5 + 2
6	2 + 5	5 + 3	18	5 + 2	6 + 4
7	3 + 5	6 + 3	19	8 + 2	8 + 2
8	6 + 3	6 + 2	20	2 + 2	3 + 3
9	8 + 3	2 + 6	21	3 + 3	2 + 2
10	2 + 4	2 + 4	22	4 + 4	4 + 4
11	5 + 3	2 + 5	23	5 + 5	5 + 5
12	2 + 6	4 + 3			

Die leichtesten (siehe Fußnote S. 13) Aufgaben stimmen gut überein, während mit ansteigender Schw die jeweiligen Rangplätze zum Teil beachtlich verschoben sind, z. B. bei 8 + 3, was wohl mit der Abhängigkeit Jugendlicher von der jeweiligen Einübung der schwierigeren Aufgaben zusammenhängt. Die Zeitwerte aus den Versuchen Brandenburgers (3) können hier nicht mitverglichen werden, da dort zweistellige Zahlen als Summanden verwendet wurden, „weil das Rechnen mit einstelligen Zahlen sehr mechanisch verläuft" (S. 76).

Die Zeitrangkurve setzt sich zusammen aus den Werten der einfachen tachistoskopischen Zeitmeßversuche mit 5 Vpn und den Ergebnissen der reaktionstachistoskopischen Zeitmeßversuche mit 3 Vpn. Der Übereinstimmungsgrad der 3 Rangordnungen der reaktionstachistoskopischen Versuche mit den 3 Rangordnungen der einfachen tachistoskopischen Methode für dieselben 3 Vpn ist ziemlich hoch (Korrelationskoeffizient für alle 3 Vpn im Durchschnitt 0,80). Um die Zusammenhänge noch weiter zu klären, werden die mittleren Variationen (in %) miteinander verglichen. Die durchschnittliche m V in % für 1 Vp beträgt bei den Rechenaufgaben (mit einfachem Tachistoskop) ± 7,8 %, für die Versuche mit dem Reaktionstachistoskop ± 9,8 %, also eine um 26 % erhöhte Streuung für die letzteren Versuche. Es ist nicht genau festzustellen, wieweit diese Erhöhung, die mit der Verfeinerung der Methode zusammenhängt, ein Ausdruck für die Sicherheit der

mechanischen Zeitstoppung im Rahmen der hier durchgeführten Versuche ist. Für die Vereinigung der Werte aus den einfachen und den reaktionstachistoskopischen Zeitmeßversuchen spricht aber neben dem hohen Übereinstimmungsgrad vor allem die dadurch erzielten günstigeren Voraussetzungen für die statistische Auswertung. Die oben näher beschriebene Zeitrangordnung (S. 20) wird also als grundlegende Rangordnung der Zeitmeßmethode betrachtet. Trotzdem können bei der weiteren Auswertung, die der hier durchgeführten Methode entsprechend vor allem auch auf individuelle Unterschiede einzugehen hat, nur die Ergebnisse der einfachen tachistoskopischen Versuche verwertet werden, da nur diese für alle 5 Vpn gleichmäßig durchgeführt wurden.

Bei einem Vergleich der 5 Rangordnungen miteinander ergeben sich beachtliche Unterschiede. Abgesehen davon, daß die Summen der absoluten Rechenzeiten und damit die auf eine Aufgabe durchschnittlich verwendeten Zeiten ziemlich schwanken (Vp A: 14, Vp B: 29, Vp C: 37, Vp D: 26, Vp E: 31 jeweils in $^1/_{100}$ sec angegeben) zeigt auch schon eine Betrachtung der Extremwerte Abweichungen voneinander. Die am stärksten bzw. am geringsten belasteten Aufgaben bei den 5 Vpn zeigt folgende Zusammenstellung:

Belastung	Vp A	Vp B	Vp C	Vp D	Vp E
Geringste Belastung ↓	3 + 2	5 + 5	5 + 5	5 + 5	2 + 2
	4 + 2	3 + 4	7 + 7	7 + 7	6 + 6
	2 + 8	4 + 3	2 + 2	3 + 3	2 + 3
	7 + 2	4 + 4	6 + 4	6 + 4	6 + 4
	2 + 2	5 + 2	4 + 4	4 + 3	3 + 2
	2 + 3	3 + 3	6 + 6	4 + 5	7 + 7
	5 + 5	6 + 6	8 + 8	8 + 2	8 + 2
↓	4 + 9	7 + 4	3 + 9	7 + 8	4 + 9
	9 + 9	5 + 8	3 + 8	4 + 7	8 + 5
	7 + 5	8 + 3	6 + 8	8 + 5	6 + 9
	4 + 7	8 + 5	4 + 7	8 + 9	3 + 8
	6 + 5	3 + 8	8 + 7	6 + 9	7 + 9
	9 + 6	6 + 9	6 + 9	6 + 8	8 + 9
Höchste Belastung	6 + 9	4 + 7	7 + 8	5 + 9	5 + 8

Auch die m V % zeigen bei den einzelnen Vpn ziemliche Unterschiede; sie betragen durchschnittlich für eine Aufgabe bei Vp A ± 6,2 %, bei Vp B ± 8,7 %, bei Vp C ± 8,3 %, bei Vp D ± 9,2 % und bei Vp E ± 10,0 %. Trotz der aufgetretenen individuellen Schwankungen ist es aber im Hinblick auf die Höhe der Korrelationskoeffizienten der Rangordnungen der 5 Vpn zur Gesamtzeitrangordnungskurve (Vp A: 0,73; Vp B: 0,91; Vp C: 0,94; Vp D: 0,92; Vp E: 0,91) doch gerechtfertigt, diese gemeinsame Rangordnung aufzustellen.

Die verhältnismäßig großen Schwankungen der m V %, die zwischen den einzelnen Vpn auftreten, lassen die Fragestellung auftauchen, ob diese unabhängig davon mit einer gewissen Regelmäßigkeit auf die einzelnen Aufgaben verteilt sind. Um einen Einblick in die Abhängigkeit der m V % von der Schw der Aufgaben zu bekommen, wurde bei jeder Vp festgestellt, wie groß die Steigerung ist, wenn die Summen der absoluten Rechenzeiten der ersten und zweiten Hälfte in jeder Rangordnung voneinander abgezogen werden; ebenso wurden die Summen der dazugehörigen m V in % verglichen. Der Vergleich der beiden Steigerungsgrade zeigt die Abhängigkeit der Streuung von der Größe der für eine Aufgabe benötigten Rechenzeit: Es ergibt sich also ein Ansteigen der Streuung mit der Schw der Aufgabe; allerdings wird diese in jedem Fall durch die Streuung schlechter angezeigt.

Eine Betrachtung der durchschnittlichen individuellen Rechenzeiten und ihre Beziehung zu den dazugehörigen Belastungsanstiegen (einstellige : zweistelligen Summen) soll zur Klärung der Frage beitragen, ob relativ schnelle Rechner auch besonders leistungsfähig sind:

Zusammenhänge bei den einzelnen Vpn	Vp A	Vp B	Vp C	Vp D	Vp E
Schw Anstieg . . .	61 %	104 %	75 %	115 %	68 %
Durchschn. Belastung in 1/100 sec	14	29	37	26	31

Es ergibt sich also die Tatsache, daß die Leistungsfähigkeit einer Person nicht durch die Höhe ihrer durchschnittlichen Reaktionsgeschwindigkeit zum Ausdruck kommt. Dadurch wird die Bedeutung der Lösungszeitdauer für eine Aufgabe als Schw Kriterium (S. 25) nicht aufgehoben, sondern nur eingeschränkt. Während dieses Kriterium in seiner grundsätzlichen Verwertbarkeit bestehen bleibt, machen die individuellen Unterschiede hinsichtlich des „Temperaments“ — abgesehen von anderen inter- und intraindividuellen Schwankungen der Leistungsfähigkeit — jede Aufstellung von allgemeingültigen Durchschnittswerten der absoluten Rechenzeiten als Ausdruck der Schw einer Aufgabe unmöglich.

Ein weiteres Schw Kriterium ergibt sich in der Verteilung der Fehlreaktionen auf die einzelnen Aufgaben bei den einzelnen Vpn. Unter Fehlreaktionen sind nicht nur die eindeutigen Falschlösungen von Aufgaben zu verstehen (z. B. 3 + 8 = 12), sondern auch jedes Stocken vor dem Aussprechen des Ergebnisses, das durch die Zeitmessung nur ungenügend erfaßt werden konnte, jeder Beginn des Versprechens (beim Aussprechen des Summen-

ergebnisses), der vom VI festgestellt werden konnte. Die folgende Tabelle gibt einen Überblick über die Verteilung der Fehlreaktionen auf die einzelnen Aufgaben bei allen 5 Vpn:

Fehlreaktionen	Dazugeh. Aufgaben	Fehlreaktionen	Dazugeh. Aufgaben	Fehlreaktionen	Dazugeh. Aufgaben
8	6 + 9	3	4 + 7	1	2 + 5
7	5 + 7		8 + 4		6 + 2
6	4 + 9		6 + 7		2 + 9
	7 + 5		7 + 8		9 + 2
	5 + 8		8 + 7		3 + 3
	8 + 5	2	2 + 2		3 + 4
5	5 + 6		3 + 9		4 + 4
	8 + 7		9 + 3		9 + 4
4	3 + 8		4 + 7		5 + 5
	7 + 6		9 + 5		5 + 6
	8 + 3		6 + 8		5 + 9
3	4 + 6		9 + 6		9 + 9

Es lassen sich unverkennbare Zusammenhänge mit der Zeitrangkurve (S. 25) feststellen. Allerdings sind unter „1" Aufgaben anzutreffen, die in der Zeitmessung sehr verschieden stark belastet sind; andererseits sind aber unter den höchstbelasteten Aufgaben (mit Fehlreaktionen) keine Aufgaben anzutreffen, die in der Zeitrangkurve auffällig gering belastet wären. Es zeigt sich also auch hier, daß die Schwierigkeit der Aufgabenlösung beeinträchtigt. Die Zahlen der Fehlreaktionen bei den einzelnen Vpn betragen: Vp A: 22, Vp B: 18, Vp C: 33, Vp D: 4, Vp E: 28; die weiblichen Vpn (Vp B und Vp D) haben die niedrigsten Zahlen von Fehlreaktionen.

Die Überprüfung der Abweichungen vom richtigen Ergebnis zeigt, daß die Abweichungen um 1 weit überwiegen. Die folgende Tabelle enthält die Abweichungen vom richtigen Ergebnis nach oben (+) und unten (—):

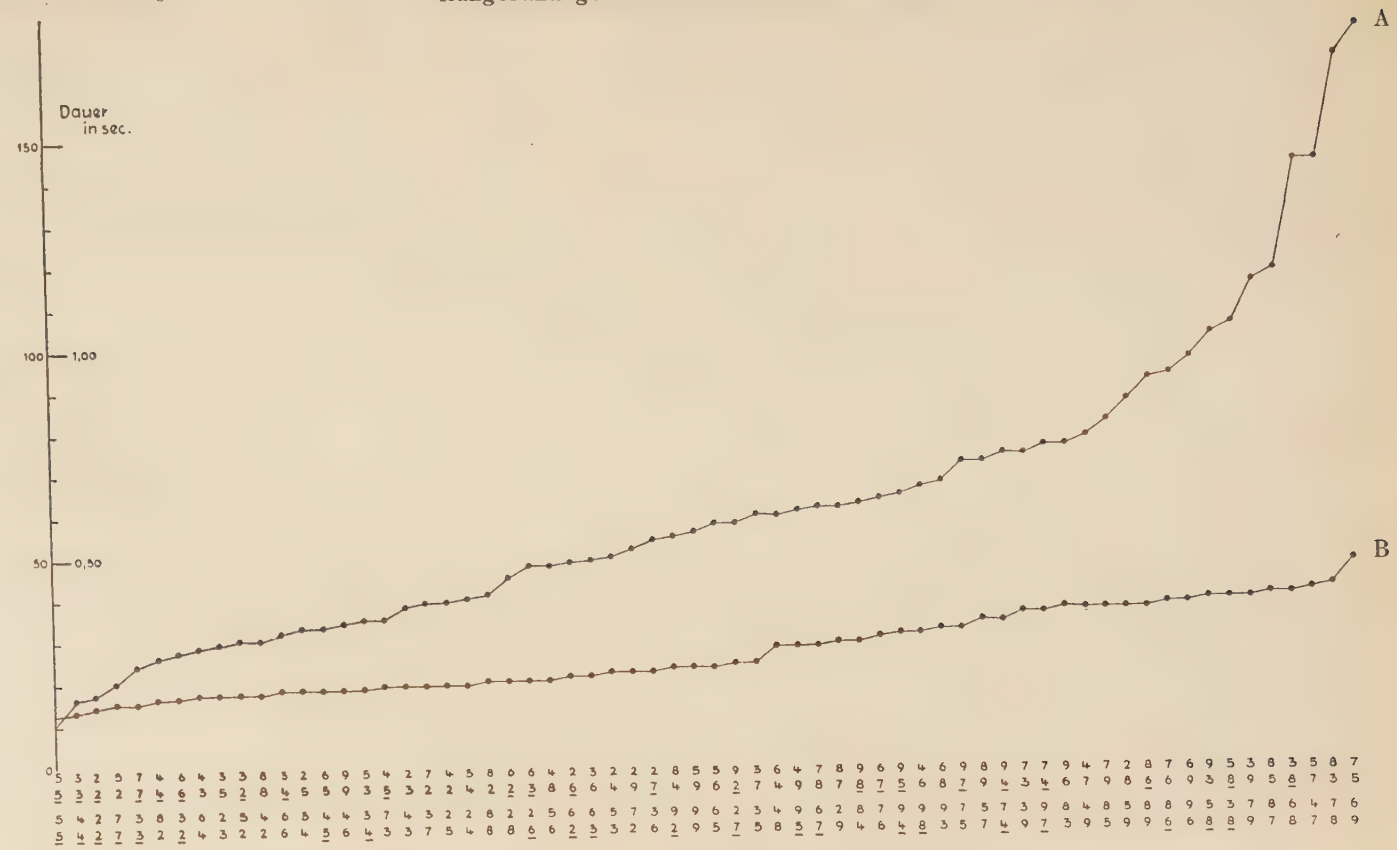
−9	−8	−7	−6	−5	−4	−3	−2	−1	+1	+2	+3	+4	+5	+6	+7	+8	+9
1	0	0	0	1	0	2	6	20	34	9	0	1	0	0	0	0	0

Die extremen Fälle verteilen sich auf die Aufgaben 9 + 6 = 6, 6 + 8 = 9 und 7 + 9 = 13. Im einzelnen kann hier auf die Arten der Fehlreaktionen nicht eingegangen werden, da, wie schon erwähnt, auch der Beginn eines Versprechens als Fehlreaktion mitgezählt wurde und sich daher nur bei einer relativ recht geringen Zahl von Fehlreaktionen eindeutig die F Art feststellen läßt.

Eine Eigentümlichkeit verdient noch Erwähnung: Alle Falschlösungen mit 5 + 7 und 7 + 5 ergaben 13, während die Falschlösungen bei 3 + 8 und 8 + 3 12 ergaben.

Zahl der Fehlleistungen

Rangordnungskurven der 1. und 2. Methode.



A: Mögliche paarweise Zahlenkombinationen geordnet nach der Belastung mit Fehlleistungen.
B: ,, ,, ,, ,, ,, ,, ,, der absoluten Rechenzeit (A M für 1 Vp in sec
(Zahlenpaare, deren Rangplätze bei beiden Methoden nicht mehr als 5 Stufen voneinander entfernt sind, sind unterstrichen.)

Vergleich der beiden Methoden.

Dem Vergleich liegen die beiden Rangordnungen, wie sie die Abbildung S. 25 zeigt, zugrunde. Dabei ergibt sich, daß die beiden Reihen in ihren 10 kleinsten Werten gut übereinstimmen, nämlich in den Aufgaben 5 + 5, 3 + 3, 2 + 2, 4 + 4, 7 + 7, 3 + 2 und 5 + 2, während sie in den obersten 10 Rangstufen nur in 3 Aufgaben übereinstimmen, nämlich in den Aufgaben 6 + 9, 5 + 8 und 3 + 8. Bei einem eingehenderen Vergleich zeigt sich jedoch deutlich der Übereinstimmungsgrad der beiden Reihen (Kk: 0,76). Da die beiden Methoden auf zwei verschiedenen Schw Kriterien beruhen, ist die auftretende Übereinstimmung auf die Schw der für beide Methoden gleichen Aufgaben zurückzuführen. Dadurch erfahren also die beiden verwendeten Methoden eine Bestätigung, d. h. sowohl die Anzahl der Fehlleistungen als auch die zur Lösung notwendige Zeitdauer ist Ausdruck der Schw einer Aufgabe.

Vor einer endgültigen Festlegung der Schw Skala für die verwendeten Aufgaben soll noch die Ergiebigkeit der verwendeten Methoden berücksichtigt werden. Die Abbildung zeigt, daß die Zeitrangkurve eine bedeutend geringere Reichweite (3,1) aufzuweisen hat als die Fehlleistungskurve (16,7). Der Höchstwert der Zeitrangkurve beträgt also nicht ganz $^1/_5$ des Höchstwertes der Fehlleistungsstatistik (aus zeichentechnischen Gründen wurde dieser Maßstab zugunsten der Zeitrangkurve auf der Abbildung etwas vergrößert.) Die 1. Methode ist also ergiebiger bei der Erfassung der Schw als die 2. Methode.

Im Hinblick auf die gefundenen individuellen Schwankungen der Zeitmeßmethode taucht die Frage auf, ob denn die 1. Methode in der angewendeten Form (einmaliger, einstündiger Versuch) genügt, um die individuelle Rangordnung genügend fein zur Ausprägung zu bringen, um damit gesicherte Unterlagen für die Gewinnung allgemeingültiger Gesetzmäßigkeiten zu haben. Es wurden zu diesem Zweck die Rangordnungen einer Vp (Vp D), die die Versuche nach der ersten und zweiten Methode mitgemacht hat, miteinander verglichen; dabei ergeben sich die erwarteten Abweichungen. In Anbetracht der hohen Zahl von Einzelversuchen bei jeder Aufgabe (14 Hauptversuche) nach der 2. Methode und der geringen Zahl von Fehlleistungen, die nach der 1. Methode auf eine Aufgabe durchschnittlich entfällt (0,23 F und 0,92 V), wird der nach der 2. Methode gefundenen Rangordnung — auch schon im Hinblick auf die viel größere Kontinuierlichkeit der Reihe — der Vorzug gegeben, d. h. die 2. Methode ermöglicht eine bessere individuelle Schw Bestimmung als die 1. Methode (jeweils in den durchgeführten Anordnungen). Grundsätzlich sind zwar allgemeingültige Schw Verhältnisse an einer geringen Anzahl von Vpn auch zu

finden; die Abgrenzung der individuellen Schwankungsbreite macht aber eine große Anzahl von Vpn notwendig. Wenn diese Abgrenzung dann einigermaßen zu Recht bestehen soll, muß vorher die individuelle Eigenart bis zu einem möglichst hohen Grad zur Ausprägung kommen. Es wäre von Interesse in weiteren Versuchen festzustellen, welcher Grad der Übereinstimmung zwischen 1. und 2. Methode sich für eine individuelle Auswertung durch eine Verfeinerung der 1. Methode erreichen ließe, etwa durch wiederholte einstündige Versuche oder durch eine weitere Steigerung der Rechengeschwindigkeit. Eine Verbesserung der 2. Methode ließe sich wohl kaum durch eine weitere Verfeinerung der Zeitmessung, wohl aber durch eine Häufung der Versuche erreichen.

Trotz der deutlichen Übereinstimmung der Rangordnungen der 1. und 2. Methode muß im Hinblick auf die großen individuellen Schwankungen nach der 2. Methode davon abgesehen werden, eine gemeinsame Rangordnung aus den Gesamtergebnissen der 1. und 2. Methode aufzustellen; vielmehr wird versucht eine Schw Abstufung durch eine Gruppeneinteilung der Aufgaben zu erreichen.

Wirth hat in seiner Arbeit (46) die Werte der Fehlleistungsstatistik und der Zeitmessung vereinigt und als Kriterium seiner Schw Rangordnung die „Fehlerdauer" verwendet; da im Rahmen vorliegender Untersuchung eine Gruppeneinteilung der Aufgaben vorgenommen wird, liegt kein Grund vor, die Werte zu vereinigen.

Einteilung der Rangordnungen (je 64 Aufgaben) in Gruppen:

2 Gruppen [50 % + 50 %]		3 Gruppen [25 % + 50 + 25 %]	8 Gruppen [8 × 12½ %]
5+5; 2+2; 5+2; 3+3; 7+7; 4+3; 3+2; 4+2; 4+4; 2+5; 6+2; 2+6; 5+4; 3+6; 2+3; 3+4; 2+4.		5+5; 2+2; 4+4.	5+5.
		9+5.	
8+9; 7+8; 4+9; 5+9; 8+7; 6+8; 9+6; 9+5; 7+4; 9+7; 4+7; 9+4; 9+3; 6+9; 8+6; 5+8; 7+6; 5+7; 8+5; 3+8; 7+5; 8+3.		6+9.	

Bei dieser Zusammenstellung werden die Rangordnungen, d. h. die 64 Aufgaben der 1. Methode (F und V einzeln) und der 2. Methode (für alle 5 Vpn einzeln) zunächst in 2 Gruppen eingeteilt. Mit Ausnahme der Aufgaben mit 2 gleichen Summanden sind in der ersten Gruppe nur Aufgaben mit einstelligen Summen und in der zweiten Gruppe nur Aufgaben mit zweistel-

ligen Summen anzutreffen. Eine weitere Einteilung in kleinere Gruppen ergibt in der ersten Gruppe stets 5 + 5, 2 + 2 und 4 + 4, in der zweiten Gruppe 9 + 5, in der dritten Gruppe 6 + 9. Die weitere Einteilung in 8 Gruppen ist bereits zu fein; nur in der ersten Gruppe (mit 8 Gliedern) ist eindeutig immer 5 + 5 anzutreffen.

Wenn die Reihenfolge der Summanden nicht berücksichtigt wird, ergibt sich folgende Gruppeneinteilung:

Einteilung der Rangordnungen (je 36 Aufgaben) in Gruppen:

2 Gruppen [50 % + 50 %]	3 Gruppen [3 × 33 1/3 %]	4 Gruppen [4 × 25 %]
5+5; 2+2; 4+4; 7+7; 3+3; 2+3; 2+4; 4+5.	5+5; 2+2; 4+4	5+5; 2+2; 4+4.
4+8; 6+7; 5+9; 3+9; 4+9; 8+9; 7+8; 7+9; 5+7; 4+7; 5+8; 6+8; 3+8; 6+9.	6+9; 4+7; 6+8; 5+7.	6+9.

Mit der Einteilung in Gruppen ist die eindeutige Feststellung gemacht, daß jede Aufgabe einer folgenden Gruppe bei allen 5 Vpn (2. Methode) und nach den Ergebnissen der 1. Methode schwieriger zu lösen ist als jede Aufgabe einer vorhergehenden Gruppe. Die Reihenfolge der Aufgaben innerhalb der einzelnen Gruppen (jeweils von links nach rechts) ist geordnet nach den Rangordnungen bei allen 5 Vpn und bei der 1. Methode. Es ist anzunehmen, daß bei einer Ergänzung des Materials nach der 2. oder einer verfeinerten 1. Methode vor allem die Rangordnungsplätze, die sich um die Gruppenübergänge verteilen, einer Korrektur ausgesetzt sein werden. Es taucht dabei die Frage auf, ob durch Vpn, die eine große Übung haben (vgl Vp A), nicht eine Schw Einteilung, die Anspruch auf Allgemeingültigkeit macht, ungültig wird. Bei genauerer Betrachtung der Versuchsergebnisse gerade der Vp A ergibt sich, daß die Frage der Grenzabsteckung, die durch die Wirksamkeit des Übungsfaktors sich notwendig zu ergeben scheint, nicht brennend wird, d. h. es ergeben sich trotz großer Übung immer noch deutliche Schw Unterschiede. Im Hinblick auf die hier untersuchten einfachen Additionsleistungen muß allerdings gesagt werden, daß die Untersuchung an noch geringer werdenden Schw Stufen nicht mehr sehr viel weiter fortgesetzt werden kann.

Im einzelnen erhalten noch verschiedene durch die 1. Methode aufgestellte Behauptungen eine Bestätigung durch die 2. Methode:

Auf Grund der 1. Methode wurde eine geringere Leistungsfähigkeit beim weiblichen Geschlecht ermittelt; dies wird durch die 2. Methode, die zwar in dieser Hinsicht bei ihrer geringen Anzahl von Vpn nicht sehr stichhaltige Ergebnisse liefern kann, bestätigt (Tabelle S. 23: Vp B und Vp D).

Lipmann (21, S. 162): „Fassen wir die Leistungen bei sämtlichen in unseren Prüfungen verwendeten Aufgabengruppen zusammen zu einem Gesamtleistungswert, so finden wir bei einem Vergleich der durchschnittlichen Leistungen von Knaben und Mädchen eine geringe männliche Überlegenheit." Pauli (31, S. 147): „Die weibliche Durchschnittsleistung bleibt in allen Fällen merklich hinter der männlichen zurück" (6—10 %). Bei Schmidberger (38, S. 140—163) wird u. a. festgestellt, „daß die geringe Überlegenheit der Knaben (8—14jährige Schulkinder) im Durchschnitt der erzielten Gesamtleistungswerte in ganz besonderem Maße auf einer männlichen Überlegenheit bei den schweren Aufgaben beruht. Entsprechend sind die Ergebnisse der Arbeiten von Fauth (9), Lindbeck (18) und Schriever (39), die beiden letzten vor allem im Hinblick auf die Vergrößerung der Unterschiede durch Übung.

Die durch die 1. Methode gefundene Gesetzmäßigkeit, daß keine Abhängigkeit der Schw von der absoluten Belastung mit Aufgaben besteht, kann mit der in der 2. Methode gefundenen Tatsache in Verbindung gebracht werden, daß die Schw Abhängigkeit einer Person nicht durch ihre durchschnittliche Rechengeschwindigkeit zum Ausdruck kommt. In beiden Fällen findet die damit von Korn (13) aufgestellte Behauptung, daß die schnellen Rechner zugleich die sicheren seien, hiedurch keine Bestätigung.

Das Ergebnis der 1. Methode, daß die Streuung mit der Schw der Aufgabe steigt, wird durch die Werte der 2. Methode bestätigt: Die m V in % beträgt für die Lesezeiten ± 2,1%, für die Rechenzeiten ± 7,8% bei den handgestoppten Versuchen (Reaktionstachistoskop: ± 2,7% und ± 9,8%).

Die Rechenzeitmessungen Brandenburgers an Jugendlichen bestätigen diese Tatsache (3, S. 80): „So ist im großen und ganzen ein Parallelgehen von Reaktionszeit und Streuung festzustellen."

Die Richtigkeit des Schw Kriteriums der 1. Methode wird auch durch die 2. Methode zum Ausdruck gebracht, insoferne auch dort die Schw die Sicherheit der Aufgabenlösung beeinträchtigt (S. 23).

Auch die Verteilung der Fehlreaktionen unterliegt in der 1. und 2. Methode denselben Regelmäßigkeiten. Neben der Erhärtung der Tatsache, daß die Abweichungen um + und — 1, und dabei besonders die + Werte, die übrigen weit überwiegen, kann auch das allgemeinere Ergebnis aufrechterhalten werden, daß die Zahl der Abweichungen mit ihrer Größe abnimmt.

Diese Ergebnisse stimmen gut überein mit den von Pauli und Korn gefundenen Gesetzmäßigkeiten: Bei Korn (13) betragen die vom richtigen Wert um 1 Einheit abweichenden Falschlösungen 59 % und davon liegen

etwa $^2/_3$ um 1 Einheit über dem richtigen Wert, und Pauli (28, S. 198) stellte fest: „Je weiter die Falschlösungen von der richtigen Lösung entfernt sind, desto seltener treten sie auf.“ Wirth (46, S. 274) findet dieselbe Regelmäßigkeiten an einfachen Subtraktionen: „Je größer die Abweichung ist, desto seltener ist sie: so daß ± 1 und ± 2 fast alles umspannen (35 % und 50 %).

Die Tasache, daß die weiblichen Vpn die niedrigsten Zahlen von Fehlreaktionen haben, steht anscheinend mit dem Ergebnis der 1. Methode in Zusammenhang, daß dort bei weiblichen Vpn die F Zahlen zugunsten der V Zahlen stark zurückgedrängt sind, was wohl auf die Tatsache zurückzuführen ist, daß die weiblichen Vpn mit einer stärker wirksamen Selbstkontrolle und mit mehr Sorgfalt im Kleinen arbeiten.

Dies kommt auch in der Feststellung Paulis (31, S. 153) zum Ausdruck: „Bei den weiblichen Leistungen zeigt sich die Neigung, die Qualität auf Kosten der Quantität zu bevorzugen, wenigstens bei dem gleichförmigen schriftlichen Addieren.“ Ein Überwiegen der Leistungsfähigkeit der männlichen Vpn hinsichtlich der Quantität der in einer Zeiteinheit gelösten Aufgaben konnte allerdings bei den vorliegenden Versuchen nicht festgestellt werden (S. 13).

Schwierigkeitsfaktoren.

Die Ursachen der objektiven Schw, d. h. der Schw, die durch individuelle Schwankungen und innerhalb dieser Schwankungen durch situativ bedingte Variationen nicht hinfällig wird, müssen restlos in der Struktur des Gegenstandes, im vorliegenden Fall in der besonderen Art der Aufgabe gesucht werden. Eine Reihe von „Schw Faktoren“ fallen bei einer Untersuchung des Gegenstandes auf: Alle Aufgaben bestehen aus 2 Zahlen, die jeweils addiert werden. Die die Schw verändernden Bedingungen können also nur in der Verschiedenheit der Zahlen und in ihrem Verhältnis zueinander liegen. Die einzelne Zahl wird dabei weniger als selbständiges Individuum als vielmehr als ein Phänomen aufgefaßt, das sich in seinen Eigenschaften durch das wechselnde Verhältnis zu anderen ändert. Im einzelnen werden nun folgende Schw Faktoren untersucht:

Hinsichtlich der Zahlen:

1. Größe der Summanden (bzw. der Summen).
2. Geradheit bzw. Ungeradheit der Summanden (bzw. der Summen).

Hinsichtlich ihres Verhältnisses zueinander:

3. Reihenfolge der Summanden.
4. Zehnerübergang.
5. Größenabstand der beiden Summanden.

Bei der folgenden Untersuchung käme es im Interesse der genauen Bestimmbarkeit darauf an, jeden einzelnen Schw Faktor isolieren zu können; das ist nicht möglich. Aus diesem Grunde sind auch Korrelationen zwischen einzelnen Schw Faktoren nicht genau durchzuführen. Es wird versucht, die Faktoren soweit zu isolieren, als dies die innere Verwobenheit ihrer Struktur zuläßt.

Neben der allgemeinen quantitativen Bestimmung der Schw Faktoren ist bei der Auswertung noch festgestellt worden, wie groß die individuelle Schwankung ist und bei welchen Aufgaben der festgestellte Schw Faktor bei allen Vpn tatsächlich auftritt; erst dann kann er eindeutig als allgemeine Schw Erscheinung erkannt werden, wenn auch die Wirksamkeit der einzelnen bekannten und unbekannten Faktoren, die bei der Lösung dieser Aufgaben sich Geltung verschaffen, Schwankungen unterworfen ist; es ist dabei nicht möglich, eindeutig festzustellen, um wieviel der Unterschied in der Wirksamkeit zweier Schw Faktoren etwa des Zehnerüberganges und der Reihenfolge der Summanden bei der Lösung z. B. der Aufgabe 3 + 8 bei einer Vp sich unterscheidet von dem einer anderen Vp.

Um einen allgemeinen Überblick über den Einfluß der Größe der Summanden bzw. ihrer Summen auf die Schw der zu lösenden Aufgabe zu geben, sind in der folgenden Zusammenstellung die durchschnittlich auf eine einstellige und auf eine zweistellige Summe entfallenden Werte zusammengestellt worden; zwecks Vergleich verschiedener Reihen wurden die Werte einander angeglichen. Die Tafel enthält also nur die relativen Summenbelastungen.

Spezifische Belastung der Summen

Summen	Zahl der		Absolute Rechenzeiten				
	F	V	Vp A	Vp B	Vp C	Vp D	Vp E
4	6,0	11,0	8,5	14,0	13,0	9,0	13,0
5	16,5	17,5	9,8	16,5	21,0	11,5	14,5
6	13,3	20,2	9,0	15,3	26,7	9,3	21,7
7	12,8	15,5	10,9	13,3	25,0	11,0	24,8
8	15,4	21,4	11,4	19,6	27,2	16,3	27,6
9	18,3	25,8	10,4	19,8	32,8	12,5	23,7
10	36,7	20,6	8,9	21,6	22,8	10,2	19,0
11	46,8	41,0	18,2	42,8	44,8	23,3	33,1
12	48,3	47,8	14,9	30,9	39,0	26,6	35,1
13	49,7	36,8	15,9	40,7	46,7	26,3	43,7
14	32,8	28,4	14,3	34,6	41,2	32,2	34,6
15	37,3	37,5	22,0	44,8	68,0	32,4	40,0
16	29,7	32,3	12,3	36,0	40,7	23,7	37,3
17	28,0	40,0	18,2	31,0	48,0	35,0	42,0
18	18,0	16,0	19,5	21,0	25,0	14,5	31,0

Die Schw Anstiege betragen (unter Berücksichtigung, daß es 6 einstellige und 9 zweistellige Summen sind) für F 165%, für V 80 %, für Vp A 61 %, für Vp B 104 %, für Vp C 75 %, für Vp D 115% und für Vp E 68%.

Das Übergewicht der F Statistik erscheint hier als ziemlich hoch; es darf in diesem Zusammenhang nicht übersehen werden, daß die Aufgaben, die 10 ergeben, nicht nur bei der Zeitmeßmethode, sondern auch bei der V Statistik niedrigere Rangplätze (im Sinne geringerer Schw) einnehmen als bei der F Statistik (S. 13). Bei einer Ausschaltung der Summen auf 10 wird das Übergewicht der F Statistik zwar vermindert, bleibt aber doch noch merklich bestehen. Das Verhältnis der einzelnen Schw Anstiege zu-

einander wird auch nicht wesentlich geändert, wenn man berücksichtigt, daß die Summen unter 10 nur 21, während die Summen über 10 36 Aufgaben in sich einschließen.

Um einen Vergleich mit den Werten von Pauli (28) und Korn (13) besser zu ermöglichen, sind noch die Verhältnisse ermittelt worden, die sich ergeben, wenn man alle Werte der einstelligen Summen zusammennimmt und allen Werten für zweistellige Summen gegenüberstellt. Der Vergleich ist allerdings nicht eindeutig durchzuführen, weil bei Pauli unter der Ordnung der einstelligen Summen auch noch andere Werte zu finden sind (Zusammenstellungen mit 1) und Korn seine Werte aus Versuchen mit Schulkindern gewonnen hat und auch hier die verglichenen Gruppen nicht ganz übereinstimmen. Bei Pauli beträgt die Schw Steigerung 220 % und bei Korn 340 %, was auch gut mit der größeren Schw Abhängigkeit der jugendlichen Vpn bei diesen Versuchen übereinstimmt. Die F Verteilung der hier durchgeführten Versuche ergibt eine Schw Steigerung von 267 %.

Auch Brandenburger (3, S. 77) fand in Übereinstimmung damit, „daß die Reaktionszeiten mit den größer werdenden Summanden steigen".

Es zeigt sich also, daß Aufgaben mit zweistelligen Summen bedeutend schwieriger zu lösen sind als solche mit einstelligen Summen.

Es wurde nun die Größe der Summanden in ihren verschiedenen Zusammenstellungen beachtet. Nach den Ergebnissen der F Statistik ergibt sich folgende Reihenfolge in aufsteigender Schw: 1. Kleiner + kleiner, 2. großer + großer, 3. kleiner + großer, 4. großer + kleiner Summand. (Kleine Summanden sind dabei 2, 3, 4, 5, große Summanden 6, 7, 8, 9.) Dieselbe Reihenfolge ergab sich bei der V Statistik. Dagegen ist bei den Zeitwerten eine andere Reihenfolge aufgetreten (Hauptrangkurve S. 25 und 4 Vpn): 1. Kleiner + kleiner, 2. großer + kleiner, 3. kleiner + großer, 4. großer + großer. Vp B hat folgende Rangordnung: 1. Kleiner + kleiner, 2. großer + kleiner, 3. großer + großer, 4. kleiner + großer; allerdings liegen hier die Werte für 3. und 4. ziemlich nahe beisammen. Eindeutig ergibt sich, daß bei der zweiten Methode stärker das gleichmäßige Größerwerden der Summanden, während bei der ersten Methode mehr der Größenunterschied der Summanden als Schw Faktor angezeigt wird. Unabhängig davon ist aber festzustellen, daß die Gruppe mit zwei kleinen Summanden bei allen Versuchen sich deutlich als die leichteste abhebt, während das Größerwerden eines oder beider Summanden jeweils eine Steigerung der Schw bedeutet.

Die Ergebnisse von Pauli decken sich gut mit den Werten der Zeitmeßmethode. Bei einem Vergleich mit der 1. Methode sind größere Unterschiede festzustellen; in den einzelnen Gruppen der 1. Methode werden nun nur die Zahlen belassen, die auch bei Pauli Verwendung fanden; eindeutig kann der Vergleich auch nicht durchgeführt werden, weil gerade die bei beiden Versuchsreihen verschieden verwendeten Aufgaben bestimmend auf die angegebenen Rangordnungen einwirken können. Es ergibt sich ein näheres Angleichen der Rangordnungen, indem sich für die hier durchgeführten Versuche eine neue Rangordnung ergibt: 1. Kleiner + kleiner, 2. großer + kleiner, 3. großer + großer, 4. kleiner + großer Summand. Wirth fand in seiner Untersuchung (46), daß sich bei Aufgaben mit längerer Zeitdauer große und ungerade Zahlen häufiger (20 %) finden als bei solchen mit kürzerer.

Um den Einfluß der Geradheit bzw. Ungeradheit der Summanden auf die Aufgabenlösung zu untersuchen, werden die Aufgaben in drei Gruppen eingeteilt; eine Aufgabe kann aus zwei geraden (1. Gruppe), zwei ungeraden (2. Gruppe) und aus einem geraden und einem ungeraden Summanden (3. Gruppe) zusammengesetzt sein. In der 1. Gruppe entfallen durchschnittlich auf eine Aufgabe 25,7 F, in der 2. Gruppe 36,1 F und in der 3. Gruppe 33,5 F; die Verteilung der V entspricht der F Verteilung: 1. Gruppe 24,5 V, 2. Gruppe 32,8 V und 3. Gruppe 32,2 V. Die Zeitmeßversuche ergeben bei 4 Vpn die Rangordnung 1., 2., 3. Gruppe, bei einer Vp (Vp D) sind 2. und 3. Gruppe in ihrer Rangordnung vertauscht.

Es wird noch das Verhältnis der geraden und der ungeraden Summen festgestellt; es ergibt sich als durchschnittliche Belastung bei F 28,6 : 35, bei V 28,2 : 35,6; ebenso läßt sich der Anstieg bei allen 5 Vpn einzeln feststellen (für absolute Rechenzeiten bei allen 5 Vpn zusammen: 126,8 : 162,3). Nach dieser Zusammenstellung zeigt sich eindeutig, daß die Ungeradheit erschwerend auf die Lösung von Aufgaben einwirkt. Eine genaue Übersicht über die Verhältnisse gibt folgende Tabelle:

Verteilung der F, V und absoluten Rechenzeiten auf die einzelnen Gruppen

Zahl d. F u. V bzw. abs. R.Z. (in $^1/_{100}$ sec)	1. Gruppe G + G	2. Gruppe U + U	3. Gruppe	
			G + U	U + G
F	411	578	543	528
V	392	524	534	497
Vp A	173	214	256	235
Vp B	380	447	559	437
Vp C	470	541	690	656
Vp D	275	347	366	320
Vp E	398	513	525	516

Der Unterschied der 1. Methode (Rangordnung: 1., 3., 2. Gruppe) zu den Ergebnissen von Pauli (Rangordnung 1., 2., 3. Gruppe) kann nicht bloß auf die mit verschiedenen Aufgaben zusammengesetzten Gruppen zurückgeführt werden, sondern es spielen auch individuelle Schwankungen herein, wie das Ergebnis der Vp D zeigt. Korn fand folgende Reihenfolge: 1., 3., 2. Gruppe. Es ist noch zu untersuchen, ob die von Pauli angegebene Schw Stufung innerhalb der Gruppe mit einem geraden und einem ungeraden Summanden mit den Ergebnissen der hier durchgeführten Versuche übereinstimmt. Pauli bezeichnet die Ungeradheit als betont, wenn die ungerade Zahl an erster Stelle steht; auf die 16 Aufgaben, bei denen der ungerade Summand vorausgeht, entfallen hier 582 F, während auf die 16 Aufgaben, bei denen der gerade Summand vorausgeht, 534 F treffen; dem entspricht die Verteilung der V 497 : 534 und das Verhältnis der Rechenzeiten 2200 : 2395,5. Daß dasselbe Verhältnis, d. h. Erschwerung, wenn der gerade Summand vorausgeht, bei allen 5 Vpn in Einzelauswertung festzustellen ist, findet die von Pauli angegebene Behauptung hier keine Bestätigung. Bei keiner Aufgabe spielt eindeutig das Vorangehen des ungeraden Summanden eine erschwerende Rolle; da-

gegen tritt bei den Aufgaben 2 + 7, 4 + 7 und 6 + 9 die angedeutete Erschwerung eindeutig nie auf, d. h. diese Aufgaben sind immer schwieriger zu lösen als ihre Umkehrungen 7 + 2, 7 + 4, 9 + 6 (für die V Statistik trifft dies nur für 6 + 9 zu).

Die Reihenfolge der Summanden hinsichtlich ihrer Größe kann eine dreifache sein: Es kann der kleinere oder der größere Summand vorausgehen oder beide Summanden können gleich groß sein. In der folgenden Zusammenstellung werden nun alle Aufgaben in diese 3 Gruppen eingeteilt (Rechenzeit in $^1/_{100}$ sec):

Verteilung auf die 3 Gruppen	Fehler	Verbesserungen	Zeitwerte für die 5 Vpn
1. Gr.: 1. < 2. S.	34,9	32,9	146,5
2. Gr.: 1. > 2. S.	36,1	32,6	130,8
3. Gr.: 1. = 2. S.	8,9	14,1	73,5

Es fragt sich, worauf die Widersprüche in der Rangordnung der 1. und 2. Gruppe bei den 3 Reihen zurückzuführen sind. Bevor das Ergebnis der F Statistik, die als die ergiebigste Methode erkannt wurde, eingehender untersucht wird, soll auf den geringen Unterschied zwischen den beiden Gruppen (0,3) bei den V und auch darauf hingewiesen werden, daß eine unter den 5 individuellen Rangordnungen bei der Zeitmessung (Vp A) dem Ergebnis der F Statistik entspricht. Um die Überdeckung von Schw Faktoren teilweise zu verhindern, wird der genau untersuchte Schw Faktor der Geradheit bzw. der Ungeradheit ausgeschaltet, d. h. zunächst wird geprüft, wie das Verhältnis ist bei zwei geraden bzw. zwei ungeraden Summanden. Da sich auch hierbei keine Eindeutigkeit feststellen läßt, wird die Betrachtung nochmals in zwei verschiedenen Abschnitten durchgeführt, d. h. es werden dieselben Verhältnisse untersucht bei Aufgaben, deren Summen 10 nicht übersteigen und dann bei Summen, deren Summen größer sind als 10. Bei geraden und ungeraden Zahlen, deren Summen unter 10 sind, ergibt sich keine Übereinstimmung mit dem Gesamtergebnis der F Statistik: Die Gruppenverhältnisse (1. Gruppe : 2. Gruppe) sind für gerade Zahlen 37 : 30, für ungerade Zahlen 16 : 14. Bei Summen über 10 dreht sich das Verhältnis um, d. h. das Vorausgehen des größeren Summanden ergibt sich als schwierigere Zusammenstellung. Bei geraden Zahlen ist das Verhältnis 1. Gruppe : 2. Gruppe 96 : 121, bei ungeraden 212 : 223 (bei letzteren macht eine Ausnahme 3 + 9). Von den Aufgaben, die 10 ergeben, ist bei 2 + 8 und 4 + 6 der Einfluß des erwähnten Schw Faktors festzustellen, nicht aber bei 3 + 7, d. h., die beiden ersten zählen in ihren Ergebnissen zu den Werten für Summen unter 10. Um die Größe des unter 10 wirksamen Schw Faktors genauer festzustellen, sind alle Aufgaben, deren Summen unter 10 liegen (ohne Rücksicht auf dabei hinzukommende neue Schw Faktoren) untersucht worden (ohne 2 + 2,

3 + 3 und 4 + 4). Das Ergebnis (165 : 126) bestätigt die oben gefundene Tatsache, daß der untersuchte SchwFaktor (Vorausgehen des kleineren Summanden) nur bei Summen unter 10 bzw. 9 eindeutig wirksam ist. An der Grenze liegen nach dieser Betrachtung die Aufgaben mit Summe 9, denn bei diesen ist die Wirkung des erwähnten Schw Faktors nicht mehr sicher festzustellen: 2 + 7 ist noch durch ihn beeinflußt, 3 + 6 und 4 + 5 nicht mehr.

Insgesamt ergibt sich also, daß das Zusammentreffen von 2 gleichgroßen Summanden eine Erleichterung in der Lösung der Aufgabe bedeutet,

außerdem, daß das Vorangehen des kleineren Summanden nur bei ganz kleinen (siehe oben) Summenergebnissen eine Erschwerung bedeutet.

Eindeutig wirkt bei beiden Methoden und innerhalb der 2. Methode bei jeder einzelnen Vp das Vorausgehen des kleineren Summanden bei den Aufgaben 2 + 8, 4 + 6 und 6 + 9 erschwerend.

Der erste Satz stimmt gut mit den Ergebnissen anderer Autoren überein, dagegen nicht der zweite Satz. Bei Jenson (12) und Ranschburg (36, S. 31) ergab sich diese Reihenfolge in aufsteigender Schw: 1. S. = 2. S., 1. S. > 2. S., 1. S. < 2. S. Brandenburger (3, S. 40) erwähnt in seiner Schrift, „daß bei der Addition die Stellung der Summanden von untergeordneter Bedeutung ist". Die von Lipmann festgesetzte Regel, die von Pauli (28) bestätigt wird, kann nicht in vollem Umfang anerkannt werden: „Bei Additionen sind die am leichtesten, bei denen die beiden Summanden gleich groß, diejenigen am schwersten, bei denen der erste Summand kleiner ist als der zweite; in der Mitte stehen diejenigen Aufgaben, bei denen der erste Summand der größere ist" (19).

Es wird nun der Zehnerübergang als Faktor der Erschwerung bzw. die Endsumme 10 als Faktor der Erleichterung untersucht.

Einen Überblick über die Verteilung der F bzw. der V und der absoluten Rechenzeiten gibt die folgende Aufstellung (Rechenzeit in $^1/_{100}$ sec):

Aufgaben	F	V	Z	Aufgaben	F	V	Z
2 + 4	20	30	91	3 + 5	16	13	126
2 + 6	17	32	104	3 + 7	39	21	100
2 + 8	60	28	95	3 + 9	67	49	183
4 + 2	14	25	80	5 + 3	14	21	111
6 + 2	16	29	103	7 + 3	41	34	92
8 + 2	29	12	69	9 + 3	47	57	148

In der ersten Reihe tritt die schon früher erwähnte schlechte Übereinstimmung der Aufgaben mit Endsumme 10 (S. 13) nach der F Statistik mit den Ergebnissen der V Statistik wieder in Erscheinung; da bei allen Vpn, die nach der Zeitmeßmethode geprüft sind und auch nach der Verteilung der V klar zum Ausdruck

kommt, daß z. B. 8 + 2 leichter ist als 6 + 2, so ist wohl der Einfluß des Zehners hier sehr deutlich festgestellt. Es ist auch nach den sonstigen Ergebnissen nicht zu erklären, wie eine Aufgabe, für die sonst alle Bedingungen gleich gegeben sind (zwei gerade Summanden, die sich ergebende Summe ist eine gerade Zahl, der größere Summand geht voraus, der Additionsschritt bleibt derselbe), die von 4 + 2 auf 6 + 2 beobachtete Schw Steigerung beim Weitergehen von 6 + 2 auf 8 + 2 sich nicht fortsetzen sollte. Die Methode der F Verteilung scheint darin eine Unzulänglichkeit zu besitzen, daß sie die Erleichterung der Lösung von Aufgaben, die die Endsumme 10 ergeben, nicht anzeigt.

In den Schw Steigerungen von 3 + 5 auf 3 + 7 und 3 + 9 sind beide Erscheinungen vereinigt: 3 + 7 kann unter erleichterten Bedingungen gelöst werden (Endsumme ist 10), während bei 3 + 9 der Zehnerübergang erschwerend auftritt. Bei jeder der 5 Vpn ist 3 + 7 leichter als 3 + 5. Wenn man sich bei der Methode der F bzw. V Verteilung die Schw Steigerung von 3 + 5 zu 3 + 9 in einer geraden Linie denkt, so liegt in beiden Fällen der Schw Punkt für 3 + 7 näher an 3 + 5. Diese Tatsache ist für die beiden Methoden auch bei 7 + 3 festzustellen, dessen Schw Punkt nach der vorigen Veranschaulichung näher an 5 + 3 als an 9 + 3 liegt; eine Ausnahme macht hier wieder die F statistische Methode. Es ist also damit eine weitere Bestätigung der Tatsache gegeben, daß die F Statistik zur Erfassung des Erleichterungsmomentes, das mit der Endsumme 10 gegeben ist, nicht geeignet ist. Das A M aus der durchschnittlich auf eine Aufgabe mit Endsumme 9 entfallenden Zahl von F und der durchschnittlich auf eine Aufgabe mit Endsumme 11 entfallenden Zahl liegt etwas unter der Zahl von F, die durchschnittlich eine Aufgabe mit Endsumme 10 belasten: 32,6 : 36,7 (36,7 = A M der F bei Aufgaben mit Endsumme 10). Für die V Statistik dreht sich das Verhältnis um: 33,4 : 20,8; eine noch größere Rolle spielt diese Tatsache nach den Ergebnissen der Zeitmeßmethode. Hier ist das Verhältnis 130,8 : 81,9 (A M in $^1/_{100}$ sec).

Die Tatsache des Zehnerüberganges als Faktor der Erschwerung zeigt folgende Zusammenstellung, die einen Überblick über die Verteilung der F bzw. der V und der absoluten Rechenzeiten für alle 5 Vpn gemeinsam gibt (Rechenzeit in $^1/_{100}$ sec):

Aufgaben	F	V	Z	Aufgaben	F	V	Z
3 + 4	20	14	84	4 + 3	10	18	82
3 + 6	18	31	106	4 + 5	13	22	91
3 + 8	93	52	210	4 + 7	41	38	223
4 + 3	10	18	82	3 + 4	20	12	84
6 + 3	24	24	99	5 + 4	17	23	78
8 + 3	101	69	192	7 + 4	38	29	184

Bei beiden Methoden und innerhalb der 2. Methode bei jeder einzelnen Vp kommt die Erscheinung des Zehnerübergangs als Schw Faktor klar zum Ausdruck. Denkt man sich 3 + 4 (bzw. 4 + 3), 3 + 6 (bzw. 6 + 3) und 3 + 8 (bzw. 8 + 3) in gleichen Abständen voneinander mit ihren dazugehörigen senkrecht gezeichneten F, V und Zeit-Werten angegeben, so ist der Winkel, der den Anstieg von 3 + 6 (bzw. 6 + 3) auf 3 + 8 (bzw. 8 + 3) angibt, jedesmal bedeutend größer als der Winkel, der den Anstieg von 3 + 4 (bzw. 4 + 3) auf 3 + 6 (bzw. 6 + 3) angibt. Dasselbe ergibt sich, wenn die Schritte von 4 + 3 (bzw. 3 + 4) auf 4 + 5 (bzw. 5 + 4) mit denen von 4 + 5 (bzw. 5 + 4) auf 4 + 7 (bzw. 7 + 4) verglichen werden.

Durch die Selbstbeobachtung einer Vp angeregt, daß ihr besonders die Aufgaben schwer gefallen seien, die einen geringen Größenunterschied aufzuweisen hatten (die schwierigste Aufgabe für diese Vp war 7 + 8), wurden noch die Größenunterschiede der Summanden bei den einzelnen Aufgaben geprüft; es ist schwer zu sagen, wie weit in den gebildeten 8 Abteilungen (s. u.) der Größenunterschied der Summanden der wirksame Schw Faktor oder nur ein Mittelwert einer größeren Anzahl anderer Schw Faktoren ist. Es wurden zwei Schw Gruppen gebildet; dabei wurde ermittelt, daß nach den Ergebnissen der beiden Methoden Aufgaben, bei denen die Summanden 5 Einheiten voneinander entfernt sind, stets in der 2. Schw Gruppe anzutreffen sind. (1. Schw Gruppe umfaßt die schwierigsten Aufgaben.) Es wurden zu dieser Betrachtung die Durchschnittswerte der auf die einzelnen Abteilungen entfallenden Aufgaben herangezogen. (Die 8 Abteilungen ergeben sich in folgender Weise: der Größenunterschied der Summanden beträgt 0, 1, 2, 3, 4, 5, 6, 7.) Im übrigen ist aber dieser Schw Faktor doch zu sehr individuellen Schwankungen unterworfen, als daß weitergehende allgemeingültige Gesetzmäßigkeiten aufgestellt werden könnten.

Zusammenfassung.

Um einen Beitrag zur Lösung der Frage zu liefern, bis zu welcher Feinheit objektive Schw Unterschiede aufgestellt werden dürfen, ohne daß sie durch individuelle Schw Rangordnungen wieder aufgehoben werden müssen, wurden Versuche an einfachen Additionen durchgeführt (64 Zusammenstellungen der Zahlen 2—9). Gleichzeitig wurden die beiden verwendeten Methoden (Fehlleistungsbelastung und Zeitdauer) miteinander verglichen. Im einzelnen wurden folgende Versuche durchgeführt:

1. 57 Vpn rechneten bei einstündigem fortlaufendem Addieren mit angeordneter Höchstgeschwindigkeit über 150 000 Aufgaben, die über 4000 Fehlleistungen einschließen.
2. 5 Vpn beteiligten sich an den Rechenzeitmeßversuchen mit einfacher tachistoskopischer Darbietung (n = 67 für jede Aufgabe), 3 Vpn an den Rechenzeitmessungen mit dem Reaktionsstachistoskop (n = 44 für jede Aufgabe). Entsprechende Versuche wurden zur Bestimmung der Lesezeit der verwendeten Aufgaben durchgeführt (n = 63 für Handstoppung, n = 30 für Reaktionstachistoskop).

Die Messung der Lesezeiten ergab mit großer Genauigkeit (m V % unter 3 %), daß die Zeitdauer für das Lesen zweier Zahlen eindeutig kleiner ist als die für die Addition derselben Zahlen benötigte Zeit. Die sehr geringen Unterschiede in den Lesezeiten ermöglichten die Aufstellung einer Lesezeitkonstanten.

Die Messung der Rechenzeiten ergab, abgesehen davon, daß sie eindeutig größer waren als die Lesezeiten, deutliche Unterschiede in der Belastung bei den einzelnen Aufgaben; ebenso traten auch bei der Fehlleistungsstatistik auffällige Belastungsunterschiede auf.

Der Vergleich der beiden Methoden (Korrelationskoeffizient der beiden Rangordnungen 0,76) ergibt **hinsichtlich der Schw der untersuchten Aufgaben eindeutige allgemeingültige Gesetzmäßigkeiten;** eine Abstufung der Schw wird durch **Einteilung der Aufgaben in Gruppen** erreicht (Seite 27, 28):

Bei einer Einteilung der 64 Aufgaben in 2 Schw Gruppen sind in der 1. Gruppe (geringe Schw) nur Aufgaben mit einstelligen Summen und in der 2. Gruppe (größere Schw) nur Aufgaben mit zweistelligen Summen zu finden (mit Ausnahme der Aufgaben mit zwei gleichen Summanden).

Was die individuellen Unterschiede betrifft, so sind Vpn mit einer hohen absoluten F bzw. V Zahl mehr von objektiven Schw Verhältnissen abhängig als solche mit einer geringen absoluten F bzw. V Zahl; dieser Zusammenhang besteht nicht für Vpn mit einer hohen gegenüber solchen mit einer geringen absoluten Aufgabenzahl. Entsprechend dazu ergibt die Zeitmeßmethode, daß Vpn mit einer höheren durchschnittlichen Rechenzeit von der Schw der Aufgaben nicht abhängiger sind als Vpn mit einer geringeren durchschnittlichen Rechenzeit (s. u.).

Eindeutig ergibt sich die geringere Leistungsfähigkeit beim weiblichen Geschlecht; außerdem neigt das weibliche Geschlecht dazu, bedeutend mehr V als F zu machen; beim männlichen Geschlecht ist die umgekehrte Tatsache festzustellen.

Die Methode der Fehlleistungsstatistik (1. Methode) ist ergiebiger zur Erfassung der Schw als die Zeitmeßmethode (2. Methode). Innerhalb der 1. Methode bringt die Verteilung der F die Schw besser zum Ausdruck als die Verteilung der V.

Die 1. Methode hat den Nachteil, daß sie die individuelle Schw Abhängigkeit nicht so scharf zum Ausdruck bringt wie die 2. Methode (jeweils in den angegebenen Arten der Durchführung). Eine Unzulänglichkeit der 1. Methode besteht auch darin, daß sie die Erleichterung der Lösung von Aufgaben, die die Endsumme 10 ergeben, nicht anzeigt.

Die Anwendbarkeit der 2. Methode erfährt insoferne eine Einschränkung, als die durchschnittliche absolute Rechenzeit einer Vp keinen Rückschluß auf die Abhängigkeit der Vp von der Schw der Aufgaben erlaubt (s. o.).

Bei der Verteilung der Fehlreaktionen ergibt sich nach der 1. und 2. Methode, daß die Abweichungen um + und — 1 (insbesonders die + Werte) vom richtigen Summenergebnis die übrigen weit übertreffen. Insgesamt nimmt die Zahl der Abweichungen vom richtigen Summenergebnis mit der Größe der Abweichung ab. Außerdem stellt sich bei den Abweichungen um + und — 1 vom richtigen Summenergebnis heraus, daß diese mehr durch die Schw der Aufgaben verursacht sind als die größeren Abweichungen.

Bei der Untersuchung der Schw Faktoren zeigen sich folgende Regelmäßigkeiten:

Die Schw der Aufgaben nimmt mit der Größe der Summanden bzw. der Summe zu.

Die Ungeradheit der Summanden bzw. der Summe wirkt erschwerend auf die Lösung von Aufgaben ein. Bei Aufgaben mit einem geraden und einem ungeraden Summanden bedeutet das Vorausgehen des geraden Summanden eine geringe Erhöhung der Schw.

Das Zusammentreffen von zwei gleichgroßen Summanden bewirkt eine Erleichterung der Aufgabenlösung.

Das Vorausgehen des kleineren Summanden hat eindeutig nur bei ganz kleinen Summenergebnissen (S. 35) einen Schw Anstieg zur Folge.

Die Lösung von Aufgaben, die in ihrer Endsumme 10 ergeben, erfährt eine Verminderung der Schw; entsprechend dazu bedeutet der Zehnerübergang bei der Addition einen Schw Anstieg.

Da sich bei diesen einfachen Aufgaben, die der normale erwachsene Mensch ohne besondere Schw lösen kann, eine allgemeine Abstufung der Schw erreichen ließ, wäre es von Interesse, festzustellen, wieweit sich die Aufstellung von objektiven Schw Graden auch bei andersgearteten Aufgaben oder bei einer allmählichen Steigerung in der Schw der zu lösenden Aufgaben erzielen ließe. Wenn auch im letzteren Fall die Versuchsbedingungen verwickelter werden, würden diese Versuche doch weit mehr als die hier durchgeführten — die nur als ein Ansatz dazu zu verstehen sind — einen Rückschluß auf allgemeingültige bzw. für bestimmte Menschengruppen gültige psychische Gesetzmäßigkeiten erlauben.

Lebenslauf.

Ich wurde in München im Jahre 1912 geboren. Nach Absolvierung der Rupprechtoberrealschule in München besuchte ich die Lehrerbildungsanstalt in Pasing und trat dann in den praktischen Volksschuldienst in Oberbayern über. Neben meiner Berufsarbeit besuchte ich an der Universität München Vorlesungen, besonders psychologischer, pädagogischer, philosophischer und anthropologischer Art. Allen meinen Lehrern, insbesonders aber Herrn Geheimrat Professor Dr. Aloys Fischer und Herrn Professor Dr. Richard Pauli möchte ich an dieser Stelle für die vielseitigen Anregungen meinen herzlichsten Dank aussprechen.

Quellen.

1. Ach N.: Analyse des Willens. (Abderhalden, Handbuch der biologischen Arbeitsmethoden 1935.) — 2. Beckmann H.: Die Entwicklung der Zahlleistung bei 2—6jährigen Kindern. Z. f. a. Ps. 22, 1923. — 3. Brandenburger K.: Die Zahlauffassung beim Schulkinde. Leipzig 1914. — 4. Döring M.: Zur Psychologie des kleinen Einmaleins. Z. Pd. Ps. 13, 1912. — 5. Eliasberg W.: Psychologie und Pathologie der Abstraktion. Z. f. a. Ps. Beih. 35, 1925. — 6. Erdmann B. und Dodge R.: Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage. Halle 1898. — 7. Erdmann B. und Dodge R.: Zur Erläuterung unserer tachistoskopischen Versuche. Z. f. Ps. 22. — 8. Erlacher C.: Vergleichende Untersuchungen über einfache Arbeitsleistungen an Schulkindern. Z. Pd. Ps. 1928. — 9. Fauth E.: Testuntersuchungen an Schulkindern nach der Methode des fortlaufenden Addierens. A. f. d. g. Ps. 51, 1925. — 10. Franken A.: Psychologische Nebenergebnisse einer tachistoskopischen Untersuchung von Zahlbildern. Z. Pd. Ps. — 11. Hylla E.: Analyse von Rechenfehlern. Z. Pd. Ps. 17. — 12. Jenson: Zur Einübung des Rechenpensums auf der Unterstufe. Päd. Zeitung 33, 1904. — 13. Korn G.: Rechenleistung und Rechenfehler. Z. f. a. Ps. 25. — 14. Kühnel J.: Neubau des Rechenunterrichts. Leipzig (Klinkhardt). — 15. Kraepelin E.: Die Arbeitskurve. Phil. Stud. v. Wundt 19. — 16. Kraepelin E.: Psychol. Arbeiten 1. — 17. Kretschmer E.: Medizinische Psychologie. Leipzig 1922. — 18. Lindbeck C.: Untersuchungen über die Arbeitskurve von Schulkindern bei fortlaufender schriftlicher Rechenarbeit. Z. Pd. Ps. 1933. — 19. Lipmann O.: Experimentelle Untersuchungen über das Rechnen. 2 Sammelberichte. Z. f. a. Ps. 5, 7. — 20. Lipmann O.: Abzählende Methoden. Leipzig 1921. — 21. Lipmann O.: Ps. Geschlechtsunterschiede. Beih. z. Z. f. a. Ps. 14 a und 14 b. 2. Auflage Leipzig 1924. — 22. Mager A.: Die Enge des Bewußtseins. Münchener Stud. z. Psych. und Philos. 1, herausgegeben von O. Külpe und K. Bühler. — 23. Mager A.: Neue Versuche zur Messung der Geschwindigkeit der Aufmerksamkeitswanderung. A. f. d. g. Ps. 53. — 24. Mayerhofer G. und Pauli R.: Die Dynamometerreaktion nebst Beschreibung eines Kinnschlüssels. A. d. psych. Inst. d. Univ. München 1933. — 25. Meyer: Beiträge zur Psychologie des kleinen Einmaleins. Z. Pd. Ps. 14. — 26. Öhrn A.: Experimentelle Untersuchungen zur Individualpsychologie. Kraepelins Psychologische Arbeiten 1. — 27. Pauli R.: Untersuchungen zur Methode des fortlaufenden Addierens. Z. f. a. Ps. Beiheft 29. — 28. Pauli R.: Beiträge zur Psychologie des Rechnens. A. f. d. g. Ps. 59, 1927. — 29. Pauli R.: Psychologisches Praktikum. Jena 1930. — 30. Pauli R.:

Eine Meßvorrichtung für psychische Zeiten. A. f. d. g. Ps. 53. — 31. Pauli R.: Zur Psychologie der Geschlechter. A. f. d. g. Ps. 66, 1928. — 32. Pauli R.: Beiträge zur Kenntnis der Arbeitskurve. A. f. d. g. Ps. 97, 1936. — 33. Petrescu-Heroiu: Vergleichende arbeitspsychologische Untersuchungen. A. f. d. g. Ps. 89, 1935. — 34. Ranschburg P.: Die Lese- und Rechenschwäche der Schulkinder. Berlin 1916. — 35. Ranschburg P.: Zur physiologischen und pathologischen Psychologie der elementaren Rechenarten. Z. E. Pd. 7, 9. — 36. Ranschburg P.: Die Leseschwäche (Legasthenie) und Rechenschwäche (Arithmasthenie) der Schulkinder im Lichte des Experiments. Berlin 1916. — 37. Schanoff B.: Die Vorgänge des Rechnens. Päd. Monogr. Leipzig 1911. — 38. Schmidberger G.: Über Geschlechtsunterschiede in der Rechenbegabung. Z. Pd. Ps. 33, 1932. — 39. Schriever H.: Untersuchungen über den Einfluß der Wiederholung und Übung auf Testleistungen. A. f. d. g. Ps. 49. — 40. Seemann J.: Die Rechenfehler. Fr. Manns Päd. Mag. 1931. — 41. Spielrein J.: Über ungleiche Schwierigkeit der einzelnen Rechenaufgaben. Zf. a. Ps. 14, 1919. — 42. Stern W.: Die differentielle Psychologie in ihren methodischen Grundlagen. Leipzig 1911. — 43. Weimer H.: Psychologie der Fehler. 2. Aufl. 1929. — 44. Weimer H.: Fehler und Irrtum. Z. Pd. Ps. 32, 1931. — 45. Wenzl A.: Theorie der Begabung. Leipzig 1935. — 46. Wirth R.: Schwierigkeitsbestimmungen an einfachsten Subtraktionen. A. f. d. g. Ps. 101, 1938. — 47. Wirth W.: Die psychotechnische Brauchbarkeit des Spearmanschen Koeffizienten. Psychotechnische Z. 2, 1925.

Druck von
J. Gotteswinter, München
Theatinerstraße 18